

⑫

DEMANDE DE BREVET D'INVENTION

A1

②② Date de dépôt : 02.09.97.

③⑦ Priorité :

④③ Date de mise à la disposition du public de la
demande : 05.03.99 Bulletin 99/09.

⑤⑥ Liste des documents cités dans le rapport de
recherche préliminaire : *Se reporter à la fin du
présent fascicule*

⑥⑦ Références à d'autres documents nationaux
apparentés :

⑦① Demandeur(s) : INSTITUT NATIONAL DE LA
RECHERCHE AGRONOMIQUE INRA — FR.

⑦② Inventeur(s) : CHOPIN MARIE CHRISTINE, CLIER
FLORENCE, EHRLICH S DUSKO, GAUTIER MICHEL
et SCHOULER CATHERINE.

⑦③ Titulaire(s) :

⑦④ Mandataire(s) : CABINET ORES.

⑤④ MECANISMES DE RESISTANCE AUX BACTERIOPHAGES R/M DE TYPE IC DE BACTERIES LACTIQUES.

⑤⑦ L'invention est relative à des polypeptides constituant
des sous-unités d'un mécanisme de résistance aux bacté-
riophages R/ M de type Ic, actif contre les phages des bac-
téries lactiques, et aux séquences d'acide nucléique codant
pour ces polypeptides.

— L'expression de ces polypeptides dans une bactérie lac-
tique permet d'augmenter sa résistance aux attaques des
bactériophages.

FR 2 767 831 - A1



**MECANISMES DE RESISTANCE AUX BACTERIOPHAGES R/M DE TYPE Ic DE
BACTERIES LACTIQUES.**

L'Invention est relative à des systèmes de
résistance aux bactériophages R/M de type Ic actifs sur les
5 bactériophages des bactéries lactiques.

L'obtention de bactéries lactiques résistantes à
l'attaque par les bactériophages, est d'un grand intérêt dans
le domaine des fermentations industrielles.

Dans ce but, il est généralement proposé :

- 10 - soit de sélectionner des mutants, naturels ou
obtenus par mutagenèse, résistants aux bactériophages ;
- soit d'utiliser des vecteurs portant des gènes
codant pour des mécanismes de résistance aux phages pour
transférer ceux-ci chez des souches initialement non-
15 résistantes.

Les mécanismes de résistance aux phages sont
regroupés en 3 classes principales, selon l'étape du cycle de
reproduction des bactériophages avec laquelle ils
interfèrent :

- 20 - les mécanismes regroupés sous la dénomination
"mécanismes d'interférence avec l'adsorption", retardent
l'adsorption du phage sur la bactérie ;
- les mécanismes dénommés "mécanismes
d'infection abortive" (Abi), bloquent la multiplication des
25 phages, et provoquent la mort de la bactérie infectée avant
que celle-ci ne produise des particules virales capables
d'infecter d'autres cellules ;
- les mécanismes dénommés "mécanismes de
restriction/modification" (R/M), dégradent l'ADN du phage dès
30 son entrée dans la bactérie. Ces mécanismes associent une
endonucléase de restriction, et une méthylase, qui a pour
fonction de protéger le génome bactérien, en modifiant les
séquences de celui-ci qui pourraient constituer des séquences
cible de l'endonucléase.

35 3 types principaux de mécanismes R/M (types I,
II, et III) ont été décrits [pour revue, cf. par exemple
BICKLE et KRUGER, Microbiological Reviews, 57, pp. 434-450
(1993)]. Les caractéristiques des mécanismes R/M des types I
et II sont brièvement rappelées ci-après :

Les mécanismes R/M de type II sont les plus fréquents ; ils sont constitués de deux enzymes distinctes, une méthylase et une endonucléase, qui sont actives séparément, et qui reconnaissent une séquence-cible commune.

5 Les mécanismes R/M de type I ont été principalement découverts chez des entérobactéries, puis ont également été mis en évidence chez quelques bactéries Gram⁺ (*Bacillus subtilis*, et *Mycobacterium pulmonis*), et chez des archaebactéries.

10 Dans ces mécanismes, la méthylase (sous-unité M, ou HsdM) et l'endonucléase (sous-unité R, ou HsdR) forment chacune une sous-unité d'un complexe enzymatique multifonctionnel où elles sont associées avec une troisième protéine (sous-unité S, ou HsdS), qui est responsable de la
15 reconnaissance de séquences spécifiques par le complexe enzymatique.

Ces trois protéines sont les produits de gènes respectivement dénommés *hsdM*, *hsdR* et *hsdS*. Des expérimentations de complémentation génétique, et la
20 comparaison des séquences de ces gènes ont permis de répartir les mécanismes R/M de type I en 4 familles distinctes n'ayant qu'une très faible homologie entre elles : Ia, Ib, Ic, et Id.

D'autre part, il a été observé que les mécanismes R/M de type I pouvaient acquérir une nouvelle spécificité à
25 la suite de recombinaisons dans le gène *hsdS* [FULLER-PACE et MURRAY, Proc. Natl. Acad. Sci. USA, 83, pp. 9368-9372, (1986) ; SHARP et al., Proc. Natl. Acad. Sci. USA, 89, pp. 9836-9840 (1992) ; GUBLER et al. EMBO J., 11, pp. 233-240, (1992)].

30 Les mécanismes R/M caractérisés jusqu'à présent chez les bactéries lactiques, et dont l'utilisation a été proposée pour rendre celles-ci plus résistantes aux bactériophages sont des mécanismes R/M de type II [FITZGERALD et al. Nucleic Acids Research, 10, pp. 8171-8179, (1982) ;
35 NYENGAARD et al., Gene, 136, pp. 371-372 (1993) ; TOWNEY et al. Gene, 136, pp. 205-209, (1993) ; DAVIS et al., Appl. Environ. Microbiol., 59, pp. 777-785, (1995) ; NYENGAARD et al., Gene, 157, pp. 13-18 (1995) ; O'SULLIVAN et al., J.

Bactériol., 177, pp. 134-143 (1995) ; MOINEAU et al., Appl. Environ. Microbiol., 61, pp. 2193-2202, (1995)].

Un des inconvénients principaux de l'utilisation de ce type de mécanismes est l'apparition de phages résistants, due en particulier au fait que la méthylase du système R/M reconnaît et modifie une certaine proportion des molécules de l'ADN des phages, qui peuvent ensuite échapper à l'action de l'endonucléase. Ce phénomène apparaît en particulier lors d'utilisation prolongée en conditions de fermentation industrielle. Pour le limiter, il a été proposé d'associer plusieurs mécanismes R/M reconnaissant des séquences-cible variées [JOSEPHSEN et KLAENHAMMER, Plasmid, 23, 71-75 (1989)] ; il est donc particulièrement souhaitable d'isoler de nouveaux mécanismes, de spécificité différente de celle des mécanismes déjà connus.

Les Inventeurs ont maintenant isolé de nouveaux mécanismes R/M de type Ic, actifs chez des bactéries lactiques, et ont exprimé les gènes codant pour leurs différents constituants.

La présente invention a pour objet un polypeptide constituant l'une des sous-unités HsdR, HsdM, ou HsdS d'un mécanisme de résistance aux bactériophages R/M de type Ic, caractérisé en ce que ledit mécanisme est actif contre les phages des bactéries lactiques.

Selon un mode de réalisation préféré d'un polypeptide constituant une sous-unité HsdR conforme à la présente invention, il comprend la séquence (I) suivante (représentée en code 1-lettre):

KRLARK

Selon un mode de réalisation préféré d'un polypeptide constituant une sous-unité HsdM conforme à la présente invention, il comprend au moins l'une des séquences suivantes (représentées en code 1-lettre) :

- la séquence (II):

GLX₁FYKYL

dans laquelle X₁ représente I ou L,

- la séquence (III):

ENDX₂NLNIPRYVDTFEEEE

dans laquelle X₂ représente Y ou F

Selon un mode de réalisation préféré d'un polypeptide constituant une sous-unité HsdS conforme à la présente invention, il comprend :

- * un domaine N-terminal d'environ 150 à 180 acides aminés, comprenant la séquence (IV) suivante :

PX₃LRFX₄GFTX₅DDWEERKX₆

dans laquelle X₃ représente E ou Q, X₄ représente E, P, D, ou K, X₅ représente N ou D, X₆ représente L ou F ;

- * un domaine central, d'environ 50 à 80 acides aminés, comprenant la séquence (V) suivante :

EQX₇KIGX₈FFKX₉LDX₁₀TX₁₁LHQRKLDLLKEQKGYX₁₂QKMFKNX₁₃KX₁₄PELRFAX₁₅FADDWEX₁₆RKLG

- dans laquelle X₇ représente R ou Q, X₈ représente S, N, ou L, X₉ représente E, H, ou Q, X₁₀ représente A, D, ou N, X₁₁ représente A, ou V, X₁₂ représente L ou F, X₁₃ représente A ou S, X₁₄ représente I ou V, X₁₅ représente E ou G, X₁₆ représente D ou E ;

- * un domaine C-terminal, d'environ 160 à 200 acides aminés, comprenant la séquence (VI) suivante :

EQX₁₅X₁₆IGSFFKQLDX₁₇TX₁₈LHQRKLX₁₉

- dans laquelle X₁₅ représente K ou Q, X₁₆ représente K ou Q, X₁₇ représente N ou D, X₁₈ représente T, A, ou V, X₁₉ représente A, ou D ;

et/ou la séquence (VII) suivante :

GFLQKMF₂₀

- dans laquelle X₂₀ représente V ou H.

La présente invention englobe en particulier :

- les polypeptides HsdR répondant à l'une des séquences respectivement représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO:2 et

SEQ ID NO:4 ;

- les polypeptides HsdM répondant à l'une des séquences respectivement représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO:6 , et SEQ ID NO:8 ;

- les polypeptides HsdS répondant à l'une des séquences respectivement représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16.

L'association de l'un quelconque des polypeptides HsdR, avec l'un quelconque des polypeptides HsdM et l'un quelconque des polypeptides HsdS conformes à l'invention permet d'obtenir un complexe enzymatique constituant un
5 mécanisme R/M de type Ic actif chez les bactéries lactiques. La spécificité d'action de ce mécanisme est conditionnée par le polypeptide HsdS. Les séquences (IV), (V), (VI) et (VII) représentent des régions conservées des polypeptides HsdS chez les bactéries lactiques (et en particulier chez les
10 lactocoques), qui sont probablement impliquées dans l'association de ces polypeptides avec les polypeptides HsdR et HsdM ; ces régions conservées sont séparées par des régions variables impliquées dans la reconnaissance de séquences nucléotidiques spécifiques.

15 La présente invention a également pour objet les séquences d'acide nucléique codant pour les polypeptides définis ci-dessus, ainsi que leurs complémentaires.

Des séquences d'ADN conformes à l'invention sont par exemple représentées par :

20 - les séquences *hsdR* SEQ ID NO:1, et SEQ ID NO:3, qui codent respectivement pour les polypeptides SEQ ID NO:2, et SEQ ID NO:4;

- les séquences *hsdM* SEQ ID NO:5, et SEQ ID NO:7, qui codent respectivement pour les polypeptides SEQ ID NO:6, et SEQ ID NO:8 ;
25

- les séquences *hsdS* SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, et SEQ ID NO:15 qui codent respectivement pour les polypeptides SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 et SEQ ID NO:16 .

30 La présente invention a également pour objet des fragments d'acide nucléique d'au moins 18 pb, homologues ou complémentaires de tout ou partie d'une séquence d'acide nucléique codant pour un polypeptide HsdR, HsdM, ou HsdS conforme à l'invention.

35 Ces fragments peuvent en particulier être utilisés comme sondes d'hybridation, et/ou amorces d'amplification, pour détecter et sélectionner des souches de bactéries lactiques, ou des plasmides hébergés par ces souches, contenant au moins une séquence codant pour un

polypeptide HsdR, HsdM, ou HsdS conforme à l'invention, et pour isoler et/ou cloner ladite séquence à partir d'une souche ou d'un plasmide sélectionné de la sorte.

Des sous-fragments préférés sont ceux qui sont
5 situés dans des séquences codant pour des régions conservées entre les sous-unités HsdR, entre les sous-unités HsdM ou entre les sous-unités HsdS des mécanismes R/M de type Ic conformes à l'invention.

Par exemple :

10 - pour sélectionner des souches de bactéries lactiques contenant une séquence codant pour une sous-unité HsdR, et/ou pour cloner ladite séquence, on utilisera avantageusement au moins un oligonucléotide homologue ou complémentaire d'une séquence codant pour au moins 6 acides
15 aminés consécutifs de l'un des peptides suivants (représentés en code 1-lettre):

TGSGKT

KRLARK

LLTGFD

20 qui correspondent à des régions conservées des sous-unités HsdR chez *L. lactis*, identifiées par les Inventeurs à partir des séquences SEQ ID NO:2 et SEQ ID NO:4

- pour sélectionner des souches de bactéries lactiques contenant une séquence codant pour une sous-unité
25 HsdM, et/ou pour cloner ladite séquence, on utilisera avantageusement au moins un oligonucléotide homologue ou complémentaire d'une séquence codant pour au moins 6 acides aminés consécutifs de l'un des peptides suivants :

FYKYLS

30 LARMNL

LPHGVLFGRGAAE

NLNIPRYVDTFEEEE

qui correspondent à des régions conservées des sous-unités HsdM chez *L. lactis*, identifiées par les
35 Inventeurs en alignant les séquences SEQ ID NO:6 et SEQ ID NO:8 ;

- pour sélectionner des souches de bactéries lactiques contenant une séquence codant pour une sous-unité HsdS, et/ou pour isoler et cloner ladite séquence, on

utilisera avantageusement au moins un oligonucléotide homologue ou complémentaire d'une séquence codant pour au moins 6 acides aminés consécutifs de l'un des peptides suivants :

5 DDWEERK
LHQRKLDLLKEQKKGY
QKMFPKNG
PELRFA
FADDWE
10 IGSFFKQLD
GFLQKMF

qui correspondent à des régions conservées des sous-unités HsdS chez *L. lactis*, identifiées par les Inventeurs à partir des séquences SEQ ID NO:10, SEQ ID NO:12, 15 SEQ ID NO:14 et SEQ ID NO:16.

Les séquences d'ADN conformes à l'invention peuvent être utilisées pour exprimer dans une bactérie lactique un ou plusieurs mécanismes R/M de type Ic, afin d'augmenter sa résistance aux bactériophages.

20 Conformément à l'invention, pour permettre l'expression dans une bactérie, en particulier une bactérie lactique, d'au moins un mécanisme de résistance aux bactériophages R/M de type Ic de bactérie lactique, on procède à la transformation de ladite bactérie avec au moins 25 une séquence d'ADN conforme à l'invention.

Le procédé conforme à l'invention peut être mis en œuvre de différentes manières. Par exemple, si la bactérie-hôte ne contient naturellement aucune séquence codant pour l'une des sous-unités HsdR, HsdM ou HsdS, elle 30 devra être transformée par au moins trois séquences d'ADN, dont chacune code pour l'un de ces polypeptides. Si, en revanche, la bactérie-hôte contient déjà au moins une séquence (portée par le chromosome bactérien ou par un plasmide, codant pour l'une des sous-unités HsdR, HsdM, ou 35 HsdS, il suffira de la transformer avec une ou des séquence(s) codant pour la ou les sous-unité(s) manquante(s).

Avantageusement, à partir d'une même souche de bactérie-hôte qui contient au moins une séquence codant pour une sous-unité HsdR et au moins une séquence codant pour une

sous-unité HsdM, on peut, par transformation avec différentes séquences codant pour des sous-unités HsdS de spécificité différentes, obtenir des souches présentant des caractéristiques de résistance différentes, et limiter
5 l'émergence de phages échappant au mécanisme de résistance.

Avantageusement, pour élargir le spectre de résistance aux bactériophages, on pourra utiliser simultanément, dans une même bactérie-hôte des séquences codant pour des sous-unités HsdS de spécificité différentes ;
10 on pourra également, dans le même but, exprimer dans la bactérie transformée un mécanisme de résistance aux bactériophages choisi parmi les mécanismes Abi, et les mécanismes R/M de type II. Dans ce cas,, il est possible d'utiliser une bactérie-hôte comprenant déjà au moins une
15 séquence codant pour un mécanisme Abi et/ou au moins une séquence codant pour un mécanisme R/M de type II, ou d'introduire à la fois dans la bactérie-hôte, la ou les séquences codant pour le(s) mécanisme(s) Abi et/ou R/M de type II, et la ou les séquence(s) codant pour la ou les sous-
20 unité(s) HsdR, HsdM ou HsdS.

Dans des bactéries transformées obtenues conformément à l'invention, les séquences codant pour les mécanismes de résistance aux bactériophages et en particulier celles codant pour les sous-unités HsdR, HsdM, ou HsdS
25 conformes à l'invention, peuvent être portées par une même molécule d'ADN (chromosome bactérien ou plasmide), ou par des molécules d'ADN différentes. Plusieurs séquences portées par une même molécule d'ADN, peuvent faire partie d'une même

unité de transcription, ou bien d'unités de transcription
30 différentes. Elles peuvent être placées sous contrôle de leur propres séquences de régulation de la transcription, ou sous contrôle de séquences hétérologues, actives dans la bactérie-hôte.

On peut par exemple exprimer les séquences codant
35 pour les sous-unités HsdR, HsdM, ou HsdS à sous contrôle de séquences promoteur permettant une expression constitutive, ou, avantageusement, sous contrôle de séquences promoteur permettant une expression inductible, par exemple un promoteur inductible par le froid tel que celui décrit dans

la Demande FR 9615731 au nom de l'INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE.

Pour mettre en œuvre le procédé selon l'invention, on peut utiliser des plasmides naturels, 5 préalablement sélectionnés à l'aide de sondes d'acide nucléique conformes à l'invention, et contenant au moins une séquence codant pour une sous-unité HsdR, HsdM, ou HsdS d'un mécanisme R/M de type Ic de bactérie lactique.

La présente invention a également pour objet des 10 vecteurs recombinants, caractérisés en ce qu'ils résultent de l'insertion d'au moins une séquence d'acide nucléique conforme à l'invention dans un vecteur approprié, et en particulier des vecteurs recombinants utilisables pour transformer des bactéries conformément à l'invention.

15 Selon l'utilisation envisagée, on peut choisir soit des vecteurs capables de se répliquer et de se maintenir dans la bactérie-hôte sous forme de plasmide, tel que par exemple les plasmides pIL252 et pIL253 décrits par SIMON et CHOPIN, [Biochimie, 70, p. 559-566, (1988)], soit des 20 vecteurs permettant l'intégration des inserts qu'il portent dans l'ADN chromosomique de la bactérie-hôte, tels que par exemple les vecteurs décrits dans la Demande PCT WO 94/16086 au nom de BIOTEKNOLOGISK INSTITUT et CHR. HASSEN'S LABORATORIUM DANMARK A/S, ou bien le plasmide pG⁺host5 décrit 25 par BISWAS et al. [J. Bact., 175, p. 3628-3635, (1995)].

De manière plus générale, des vecteurs plasmidiques ainsi que des vecteurs d'intégration utilisables chez les lactocoques sont décrits dans la revue de : 30 LEENHOUTS K.J. and VENEMA G. (1993). Lactococal plasmid vectors, In : K.G. HARDY (ed) Plasmids. A practical approach. Second Edition. IRL Press, Oxford, p. 65-94.

Si plusieurs vecteurs sont utilisés pour introduire dans la bactérie-hôte les différentes sous-unités HsdR, HsdM, et HsdS, et éventuellement d'autres mécanismes de 35 résistance aux bactériophage, on choisira des vecteurs compatibles entre eux.

La présente invention peut être avantageusement mise en œuvre dans tous les domaines où intervient une fermentation industrielle, et en particulier dans l'industrie

laitière et fromagère. Des souches bactériennes possédant une résistance accrue aux bactériophages peuvent être obtenues conformément à l'invention, à partir de souches de bactéries lactiques d'intérêt industriel, soit par sélection de
5 bactéries possédant naturellement un système R/M de type Ic grâce aux sondes d'acide nucléique conformes à l'invention, soit par transformation de bactéries-hôte à l'aide de séquences d'acide nucléique conformes à l'invention.

La présente Invention sera mieux comprise à
10 l'aide du complément de description qui va suivre, qui se réfère à des exemples de clonage de séquences codant pour des sous-unités de mécanismes R/M de type Ic, et d'utilisation de ces séquences pour accroître la résistance de bactéries hôte aux attaques des bactériophages.

15 **EXEMPLE 1 : OBTENTION DE FRAGMENTS D'ADN CODANT POUR DES SOUS-UNITES HSDS DE MECANISMES R/M DE TYPE Ic**

La manipulation de l'ADN, le clonage et la transformation de cellules bactériennes sont, en l'absence de précisions contraires, effectués selon les protocoles décrits
20 par SAMBROOK et al. [(Molecular cloning : a laboratory manual., Cold Spring Harbor Laboratory, Cold Spring Harbor N.Y.(1989)].

GAUTIER et CHOPIN [App. Environ. Microbiol. 53, 923-927 (1987)], et CHOPIN et al., [Plasmid, 11, pp.260-263,
25 (1984)] ont décrit l'existence de mécanismes de résistance aux phages de type R/M, respectivement associés aux plasmides pIL 103 et pIL7.

L'ADN de pIL 103 a été digéré par EcoRI, et les
fragments obtenus ont été ligaturés au site EcoRI du vecteur
30 pIL204 [SIMON et CHOPIN, Biochimie 70, 559-566, (1988)]. Le mélange de ligation a été utilisé pour transformer des bactéries *L. Lactis* ssp *lactis* de la souche IL1403. Les colonies bactériennes ont été sélectionnées sur la base de leur résistance au phage bIL67. Un plasmide recombinant
35 portant un insert de 3,7 kb environ a été obtenu à partir de clones bactériens résistant à l'attaque phagique. Ce plasmide recombinant a été dénommé pIL261.

De manière surprenante, le séquençage de l'insert de 3,7 kb du plasmide pIL261 n'a fait apparaître aucune séquence présentant une homologie avec celles des mécanismes R/M connus chez les lactocoques.

5 Cependant, 2 cadres de lecture *orf1* et *orf2* ont été localisés sur ce fragment de 3,7 kb ;

Le cadre de lecture *orf1* code pour une protéine présentant une forte homologie avec des protéines RepB déjà identifiées chez les lactocoques.

10 La séquence de la protéine codée par le cadre de lecture *orf2* ne présente aucune homologie avec des séquences connues de protéines de bactéries lactiques ; toutefois, la région centrale et la région N-terminale de cette protéine contiennent des séquences répétées possédant une certaine
15 homologie avec des séquences consensus répétées des sous-unités HsdS de systèmes R/M de type Ic des entérobactéries et des mycoplasmes.

Une séquence d'acide nucléique contenant le cadre de lecture *orf2*, et la séquence du polypeptide correspondant
20 sont respectivement représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO : 13 et SEQ ID NO : 14.

De manière similaire, le séquençage du plasmide pIL7 a permis de mettre en évidence un cadre de lecture codant pour une protéine présentant une homologie importante
25 avec celle codée par le cadre de lecture *orf2* du plasmide pIL261, et possédant également des séquences répétées rappelant celles des sous-unités HsdS des systèmes R/M de type I.

Une séquence d'acide nucléique contenant ce
30 troisième cadre de lecture, et la séquence du polypeptide correspondant sont respectivement représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO : 15 et SEQ ID NO : 16.

Des sondes oligonucléotidiques, dérivées des
35 séquences codant pour des régions conservées entre les séquences SEQ ID NO : 14 et SEQ ID NO : 16 ont été utilisées pour rechercher l'existence d'autres séquences homologues de type *hsdS*. Ces sondes ont permis la mise en évidence de séquences de ce type, sur l'ADN chromosomique de la souche

IL1403, ainsi que sur un plasmide issu de la souche de *L. lactis* ssp. *lactis* IL420, hébergé par IL403.

La séquence de type *hsdS* obtenue à partir de l'ADN chromosomique de la souche IL1403, et la séquence du polypeptide correspondant sont respectivement représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO : 9 et SEQ ID NO : 10

La séquence de type *hsdS* obtenue à partir du plasmide issu de la souche de *L. lactis* ssp. *lactis* IL420, et la séquence du polypeptide correspondant sont respectivement représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO : 11 et SEQ ID NO : 12.

EXEMPLE 2 : OBTENTION DE FRAGMENTS D'ADN CODANT POUR DES SOUS-UNITES HsdR ET HsdM DE MECANISMES R/M DE TYPE Ic

Les régions du chromosome de la souche IL1403, et du plasmide issu de la souche IL420 sur lesquelles ont été localisées les séquences de type *hsdS* ont été entièrement séquencées.

Ce séquençage a permis de mettre en évidence, dans les 2 cas, directement en amont de la séquence *hsdS*, 2 cadres de lecture ouverte, codant pour des protéines contenant respectivement des régions présentant une homologie importante avec des domaines conservés des sous-unités R et M des systèmes R/M de type Ic d'entérobactéries et de mycoplasmes. Ces 2 cadres de lecture ouverte ont été respectivement dénommés, du fait de cette homologie, *hsdR* et *hsdM*.

Sur le chromosome de la souche IL1403, comme sur le plasmide issu de la souche IL420, l'ordre des cadres de lecture ouverte est, d'amont en aval : *hsdR*, *hsdM* et *hsdS*.

La séquence de type *hsdR* obtenue à partir de l'ADN chromosomique de la souche IL1403, et la séquence du polypeptide correspondant sont respectivement représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO : 1 et SEQ ID NO : 2.

La séquence de type *hsdM* obtenue à partir de l'ADN chromosomique de la souche IL1403, et la séquence du polypeptide correspondant sont respectivement représentées

dans la liste de séquences en annexe sous les numéros SEQ ID NO : 5 et SEQ ID NO : 6.

La séquence de type *hsdR* obtenue à partir du plasmide issu de la souche de *L. lactis* ssp. *lactis* IL420, et la séquence du polypeptide correspondant sont respectivement représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO : 3 et SEQ ID NO : 4

La séquence de type *hsdM* obtenue à partir du plasmide issu de la souche de *L. lactis* ssp. *lactis* IL420, et la séquence du polypeptide correspondant sont respectivement représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO : 7 et SEQ ID NO : 8.

EXEMPLE 3 : UTILISATION D'OLIGONUCLEOTIDES DERIVES DES SEQUENCES HSDR, HSDM ET HSDS POUR LA DETECTION DE SOUCHES BACTERIENNES POSSEDANT DES SOUS-UNITES DE SYSTEMES R/M DE TYPE Ic.

Des régions des séquences *hsdM* du plasmide issu de la souche IL420, et du chromosome de la souche IL1403, codant pour les mêmes séquences peptidiques ont été identifiées.

Les oligonucléotides suivants :

* oligo 66 : TTA GCA CGT ATG AAC TTA

* oligo 67 : TTG GCT CGA ATG AAT TTA

représentent des séquences qui codent respectivement dans le plasmide de la souche IL420, et dans le chromosome de la souche IL1403, pour la séquence peptidique (code 1 lettre) LARMNL.

Les oligonucléotides suivants :

* oligo 68 : CTC TTC AAA GGT ATC CAC

* oligo 69 : TTC CTC AAA GGT ATC TAC

représentent des séquences complémentaires des séquences codant respectivement, dans le chromosome de la souche IL1403, et dans le plasmide de la souche IL420, pour la séquence peptidique (code 1 lettre) VDTFEE.

Ces oligonucléotides sont utilisés comme amorces d'amplification par PCR (couples 66/69 et 67/68) sur 16 souches différentes de lactocoques.

Les conditions d'amplification utilisées sont les suivantes : l'ADN extrait de chacune des souches est mis en présence de désoxyribonucléotides triphosphate (0,2 mM de chaque), de 0,1 μ M de chaque amorce, et de 2,5 Unités de Taq polymérase (PROMEGA) dans un tampon standard pour Taq polymérase (PROMEGA). La réaction s'effectue dans un volume final de 100 μ l.

Après dénaturation (94°C, 5 minutes), on effectue 30 cycles alternant une étape de fixation des amorces à 50°C pendant 30 secondes, une étape d'extension à 72°C pendant 30 secondes, et une étape de dénaturation à 94°C pendant 30 secondes.

Comme le montre le Tableau I ci-dessous, pour 8 des 16 souches testées (IL582, IL858, IL910, IL993, IL964, IL827, IL854, MG1363) on observe, avec l'un ou l'autre des 2 couples d'amorces la présence d'un produit d'amplification de la longueur attendue (environ 670 pb).

TABLEAU I

SOUCHE	AMORCES	
	66/69	67/68
IL827 (D)	+	++
IL910 (L)	++	-
IL2967 (L)	-	-
CNRZ194 (C)	-	-
CNRZ340 (D)	-	-
CNRZ106 (C)	-	-
IL2034	-	-
IL854 (C)	+/-	++
IL993 (C)	-	++
IL858 (C)	+	++
CNRZ10S (C)	-	-
CNRZ269 (D)	-	-
IL582 (L)	+/-	++
IL960 (L)	-	-
IL964 (C)	+/-	++
MG1363 (L)	+	-

(L) : *ssp. lactis*

(D) *ssp. lactis biovar. diacetylactis*

(C) *ssp. cremoris*

- 5 ++ produit d'amplification nettement détectable ;
+ produit d'amplification détectable ;
+/- produit d'amplification faiblement détectable ;
- pas de produit d'amplification détectable.

10 Les produits d'amplification obtenus à partir des souches IL582, IL858, IL910, IL993, MG1363, IL827, IL854, IL964, ont été séquencés. Chez les souches IL582, IL858, IL993, IL827, IL854, IL964, la séquence est très semblable à la séquence *hsdM* du chromosome de la souche IL1403, et chez les souches IL910 et MG1363 la séquence est très semblable à la séquence *hsdM* du plasmide de la souche IL420.

EXEMPLE 4 : RESISTANCE A L'ATTAQUE PHAGIQUE CONFEREES PAR LES SOUS-UNITES HsdR, HsdM ET HsdS.

Les souches utilisées sont : les souches MG1363, et IL582 qui possèdent chacune une séquence codant pour la

sous-unité M d'un système R/M de type Ic, et la souche IL1403, qui possède sur son chromosome des séquences codant pour les 3 sous-unités R, M et S d'un système R/M de type Ic.

Les bactéries sont cultivées et infectées par les
5 phages dans les conditions décrites par TERZAGHI et SANDINE, [Appl. Microbiol., 29, 807-815, (1975)]

Les résultats de l'infection sont exprimés en PFU/ml, et l'efficacité de propagation (eop) est déterminée, pour chaque essai, par le rapport entre le nombre de PFU/ml
10 pour la souche bactérienne testée, et le nombre de PFU/ml pour une souche bactérienne choisie comme témoin.

Le tableau II ci-dessous montre les résultats obtenus avec le phage c2 sur la souche MG1363, choisie comme témoin, et la souche IL1403.

15 TABLEAU II

SOUCHE	NOMBRE DE PHAGES (PFU/ML)	EOP
MG1363	10^9	1
IL1403	$2,8 \times 10^7$	3×10^{-2}

Le tableau III ci-dessous montre les résultats obtenus avec le phage bIL67 (préalablement propagé sur la souche IL1403) sur :

- la souche IL1403 sans plasmides (1), et
 - 20 - la souche IL582 (5)
- choisies comme témoins,
- la souche IL1403 transformée avec le plasmide issu de la souche IL420, portant des séquences *hsdR*, *hsdM* et *hsdS* (2).
-
- 25 - la souche IL1403 transformée avec le plasmide pIL7, portant une séquence *hsdS* (3)
 - la souche IL1403 transformée avec le plasmide pIL261 portant une séquence *hsdS* (4)
 - la souche IL582 transformée avec le plasmide
 - 30 pIL261 portant une séquence *hsdS* (6)

TABLEAU III

SOUCHE	NOMBRE DE PHAGES (PFU/ML)	EOP
1	6×10^9	1
2	3×10^4	5×10^{-6}
3	$3,7 \times 10^7$	6×10^{-3}
4	4×10^3	6×10^{-7}
5	4×10^9	1
6	3×10^4	$1,8 \times 10^{-5}$

Ces résultats montrent que :

- * Les sous-unités HsdS des différents mécanismes de résistance ont effectivement une spécificité différente :
- 5 le mécanisme porté par le chromosome de la souche IL1430 est faiblement efficace contre le phage bIL67 , alors qu'en revanche les sous-unités HsdS portées par les plasmides augmentent, à un degré plus ou moins élevé, la résistance à ce phage.
- 10 * Les sous-unités HsdS codées par les plasmides pIL7 et pIL261 peuvent s'associer avec les sous-unités HsdR et HsdM codées par le chromosome de IL1430 pour reconstituer un complexe R/M de type Ic fonctionnel.

LISTE DE SEQUENCES

(2) INFORMATIONS POUR LA SEQ ID NO: 1:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 3050 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: double
- (D) CONFIGURATION: linéaire

(vi) ORIGINE:

- (A) ORGANISME: Lactococcus lactis

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: CDS
- (B) EMBLEMENT: 44..3031
- (D) AUTRES INFORMATIONS: /product= "HsdR"

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:

AAAGATAACA AAAATCATGC ATTATTGAGA GGGAGAAAAA GGA ATG GCA GAA GCA	55
Met Ala Glu Ala	
1	
AAA TTT GAA GCT GCT TTG ATC AAA AAA CTA GAA GCA GAA GGA TGG ACC	103
Lys Phe Glu Ala Ala Leu Ile Lys Lys Leu Glu Ala Glu Gly Trp Thr	
5 10 15 20	
TAT CGT AAA GAT TTA TCT TAT GTC TCT ATT AAG GTC TTA GAG GGA CAC	151
Tyr Arg Lys Asp Leu Ser Tyr Val Ser Ile Lys Val Leu Glu Gly His	
25 30 35	
TGG CGT GAG GTG CTT AAT GAA AAT AAT GCT TAT AAA TTA AAT GGC AAA	199
Trp Arg Glu Val Leu Asn Glu Asn Asn Ala Tyr Lys Leu Asn Gly Lys	
40 45 50	
CCT TTG TCT GAT GTT GAA TTT GGC TTA GTG ATG CAA GAA GTT CAA AGG	247
Pro Leu Ser Asp Val Glu Phe Gly Leu Val Met Gln Glu Val Gln Arg	
55 60 65	
ATT AAA ACG CCT TAT GAT GCT CAA CTT TTA TTG GTA GGG GCA GGA GGT	295
Ile Lys Thr Pro Tyr Asp Ala Gln Leu Leu Leu Val Gly Ala Gly Gly	
70 75 80	
GTC GGC TCT ATT CCT ATA ACA CGA GAT GAC GGC TCT AAC TTA GAA GTA	343
Val Gly Ser Ile Pro Ile Thr Arg Asp Asp Gly Ser Asn Leu Glu Val	
85 90 95 100	
GAG ATT TTT TAT GAA GAT GAT GTG GCT GGT GGT CGG TCA CGC TAT GAA	391
Glu Ile Phe Tyr Glu Asp Asp Val Ala Gly Gly Arg Ser Arg Tyr Glu	
105 110 115	
GTC GTC AGC CAA GTG ATT TTT GAT GAG CTC CCT CAT GGA CTA GCG AGC	439
Val Val Ser Gln Val Ile Phe Asp Glu Leu Pro His Gly Leu Ala Ser	
120 125 130	
AAA AGA ATT ATT GAT TTA GCC TTG CTT ATT AAC GGA ATC CCA GTA GCG	487
Lys Arg Ile Ile Asp Leu Ala Leu Leu Ile Asn Gly Ile Pro Val Ala	
135 140 145	

CAT ATT GAA GAA AAA GAT GAA CAC TTA CAA AAC CAA TGG GGC GCT TTT His Ile Glu Glu Lys Asp Glu His Leu Gln Asn Gln Trp Gly Ala Phe 150 155 160	535
GAG CAA TTA AAA GGG TAT CAC GGA GAA GGC TTA TAT GAG GGA TTA TTT Glu Gln Leu Lys Gly Tyr His Gly Glu Gly Leu Tyr Glu Gly Leu Phe 165 170 175 180	583
GCT TTT GTT CAA GTT CAA TTT ATC TTG AGT CAA CAC TCT GCC AAC TAT Ala Phe Val Gln Val Gln Phe Ile Leu Ser Gln His Ser Ala Asn Tyr 185 190 195	631
TTT GCC CGT CCT AAT CGT CTT GAA AAT TAT AAC AAA ACT TTT GTT TTT Phe Ala Arg Pro Asn Arg Leu Glu Asn Tyr Asn Lys Thr Phe Val Phe 200 205 210	679
GGT TGG CGA GAT GAG CAA CAG AAA GAT ATC ACT GAT GCG TTT GTC TTT Gly Trp Arg Asp Glu Gln Gln Lys Asp Ile Thr Asp Ala Phe Val Phe 215 220 225	727
GCC CAT CAA GTG TTA GGA ATC CCT GCA CTT CAT CGA CTA GTG ACT GTG Ala His Gln Val Leu Gly Ile Pro Ala Leu His Arg Leu Val Thr Val 230 235 240	775
AAC ATG ATT CCA GAT GCT TCC AAT AGT AAT TTG ATG GTG ATG AGA AGT Asn Met Ile Pro Asp Ala Ser Asn Ser Asn Leu Met Val Met Arg Ser 245 250 255 260	823
TAT CAA ATT CAA GCC ACA AGA GCG ATT TTA CAG CGT ATG AAA GAA ATG Tyr Gln Ile Gln Ala Thr Arg Ala Ile Leu Gln Arg Met Lys Glu Met 265 270 275	871
GAA CAT AAC GAC TAC ATT CAA AAG GAA GGG GGC TAT ATC TGG CAT ACC Glu His Asn Asp Tyr Ile Gln Lys Glu Gly Gly Tyr Ile Trp His Thr 280 285 290	919
ACT GGG TCG GGT AAA ACG GTG ACT TCT TTT AAA GTA GCC CAG CTA CTA Thr Gly Ser Gly Lys Thr Val Thr Ser Phe Lys Val Ala Gln Leu Leu 295 300 305	967
GCG GCT GCC CCT AAA GTA AAA AAT GTT TTA TTT ATT GTT GAT CGT GTG Ala Ala Ala Pro Lys Val Lys Asn Val Leu Phe Ile Val Asp Arg Val 310 315 320	1015
GAC TTA GTG GAT CAA ACG TTA GAA AAC TTT AAA GAT TTT GCT TAC ATT Asp Leu Val Asp Gln Thr Leu Glu Asn Phe Lys Asp Phe Ala Tyr Ile 325 330 335 340	1063
CAA TTT AAA AAT AGA ATT AAA AAA GTA AAT GGT AGG GAG TTA AAA CGA Gln Phe Lys Asn Arg Ile Lys Lys Val Asn Gly Arg Glu Leu Lys Arg 345 350 355	1111
GAG TTG AGC CAT AAA GGT GCT TCA CAG ATC TTA TTA ATT TCT GTC CAA Glu Leu Ser His Lys Gly Ala Ser Gln Ile Leu Leu Ile Ser Val Gln 360 365 370	1159
GGT TTA ACA AAA GCC GTT AAA AAT GGG CTA AAA AAT ACC GAT CGA AAT	1207

Gly	Leu	Thr	Lys	Ala	Val	Lys	Asn	Gly	Leu	Lys	Asn	Thr	Asp	Arg	Asn		
			375				380						385				
GTC	ATT	ATT	ATG	GAT	GAA	GCA	CAT	CGA	AGT	GCT	AAT	GGA	GAA	TCG	GTT	1255	
Val	Ile	Ile	Met	Asp	Glu	Ala	His	Arg	Ser	Ala	Asn	Gly	Glu	Ser	Val		
			390				395					400					
CAG	CTT	ATT	AAA	AGT	GCC	TTT	CAA	AAA	ACG	ACT	TGG	TTT	GGG	TTT	ACA	1303	
Gln	Leu	Ile	Lys	Ser	Ala	Phe	Gln	Lys	Thr	Thr	Trp	Phe	Gly	Phe	Thr		
					405		410				415				420		
GGA	ACT	CCT	AAT	TTT	TAT	AGT	GAC	GAG	ATT	AAT	GAC	GTT	CAA	ACC	ACT	1351	
Gly	Thr	Pro	Asn	Phe	Tyr	Ser	Asp	Glu	Ile	Asn	Asp	Val	Gln	Thr	Thr		
				425					430					435			
CGA	GAC	ATT	TCA	ACC	CAT	GAT	ATT	TTT	GGG	AAG	AGA	CTC	CAT	TCT	TAT	1399	
Arg	Asp	Ile	Ser	Thr	His	Asp	Ile	Phe	Gly	Lys	Arg	Leu	His	Ser	Tyr		
			440					445						450			
ACC	ATC	AAG	GAT	GCG	ATT	GGA	GAT	GGA	AAC	GTG	TTA	GGG	TTT	GAT	ATT	1447	
Thr	Ile	Lys	Asp	Ala	Ile	Gly	Asp	Gly	Asn	Val	Leu	Gly	Phe	Asp	Ile		
			455				460						465				
ACT	TAC	TTT	AAT	CCC	ACG	ATT	GAA	ATT	GAA	TCA	CTT	GAC	GAA	GAA	CAT	1495	
Thr	Tyr	Phe	Asn	Pro	Thr	Ile	Glu	Ile	Glu	Ser	Leu	Asp	Glu	Glu	His		
			470				475					480					
TCC	GAA	AAA	GAT	TAT	GAA	AAA	GAA	GTT	TAT	CAA	AGT	CAT	GTT	TAT	CGA	1543	
Ser	Glu	Lys	Asp	Tyr	Glu	Lys	Glu	Val	Tyr	Gln	Ser	His	Val	Tyr	Arg		
			485			490				495					500		
GAA	CAA	GTG	GTG	CAA	GAC	ATT	TTA	AAC	CTT	TGG	GAT	AAA	ACG	TCA	AGT	1591	
Glu	Gln	Val	Val	Gln	Asp	Ile	Leu	Asn	Leu	Trp	Asp	Lys	Thr	Ser	Ser		
				505					510					515			
GGG	GCT	TTG	GTC	GCT	GGC	AAA	CGA	GAA	AAG	AAT	GTT	TTT	CAG	GCG	ATG	1639	
Gly	Ala	Leu	Val	Ala	Gly	Lys	Arg	Glu	Lys	Asn	Val	Phe	Gln	Ala	Met		
			520					525					530				
TTA	GCT	GTT	TCA	GGA	AAG	CAA	GCC	GTG	GTT	CAT	TAC	TAC	AAT	CTT	TTT	1687	
Leu	Ala	Val	Ser	Gly	Lys	Gln	Ala	Val	Val	His	Tyr	Tyr	Asn	Leu	Phe		
			535				540						545				
AAA	GAA	AAA	GCC	CCT	CAT	TTG	AGG	GTA	GCA	ATG	ACT	TTT	TCT	CGA	GAT	1735	
Lys	Glu	Lys	Ala	Pro	His	Leu	Arg	Val	Ala	Met	Thr	Phe	Ser	Arg	Asp		
			550				555					560					
GAA	TCT	AAT	CAA	CCT	GGA	ACA	AAA	GAA	CAG	AAT	GAA	GCC	TTA	AAA	AAA	1783	
Glu	Ser	Asn	Gln	Pro	Gly	Thr	Lys	Glu	Gln	Asn	Glu	Ala	Leu	Lys	Lys		
			565				570				575				580		
GCG	ATC	AAA	GAA	TAC	AGT	AGT	TGC	TTC	AAT	GTC	CCT	AGT	CTT	TTA	AAT	1831	
Ala	Ile	Lys	Glu	Tyr	Ser	Ser	Cys	Phe	Asn	Val	Pro	Ser	Leu	Leu	Asn		
				585					590					595			
GCT	CAA	GAA	CCG	GCA	CGG	GCC	TAC	ATG	ATT	GAT	ATC	ACT	AAA	CGG	CTC	1879	
Ala	Gln	Glu	Pro	Ala	Arg	Ala	Tyr	Met	Ile	Asp	Ile	Thr	Lys	Arg	Leu		
			600					605						610			

GCT CGT AAA AAA CCA TAT AAC CAA GGA AAA GAT GAG GAC CGA TTA GAT	1927
Ala Arg Lys Lys Pro Tyr Asn Gln Gly Lys Asp Glu Asp Arg Leu Asp	
615 620 625	
CTA GTG ATT GTC TCG GAC CAA CTG CTG ACA GGA TTT GAC TCA AAA TAT	1975
Leu Val Ile Val Ser Asp Gln Leu Leu Thr Gly Phe Asp Ser Lys Tyr	
630 635 640	
ATT AAT ACG ATT TAT ATG GAT AAA CAG CTG AGA GAA GGG ATG TTA ATC	2023
Ile Asn Thr Ile Tyr Met Asp Lys Gln Leu Arg Glu Gly Met Leu Ile	
645 650 655 660	
CAA GCG ATG TCC CGG ACC AAT CGA ACG TTT CAC TTG AAC AGT AAA CCT	2071
Gln Ala Met Ser Arg Thr Asn Arg Thr Phe His Leu Asn Ser Lys Pro	
665 670 675	
CAC GGA AAA GTT CGG TTC TAT CGT CAA GGC GAA CAA ATG AAG TCT TTT	2119
His Gly Lys Val Arg Phe Tyr Arg Gln Gly Glu Gln Met Lys Ser Phe	
680 685 690	
GTA GAA AAT GCC CTT CGG ATT TAT ACA CGA GGG GGG AAT GAT ACG CTT	2167
Val Glu Asn Ala Leu Arg Ile Tyr Thr Arg Gly Gly Asn Asp Thr Leu	
695 700 705	
CAG GGA GCA GAT GAA GAC TCC AAA AAT GAA GAT ATC CAA TCC TTA GAA	2215
Gln Gly Ala Asp Glu Asp Ser Lys Asn Glu Asp Ile Gln Ser Leu Glu	
710 715 720	
GAC GAG CAT ATC CTG GCA GAA CCC CAA AGT CAT CAA ATC CCA AAA TTA	2263
Asp Glu His Ile Leu Ala Glu Pro Gln Ser His Gln Ile Pro Lys Leu	
725 730 735 740	
ACA CCC GCT GTT CAA GAA CTA AAA GCC TTT GCG GGA GAA GAT TTT AGT	2311
Thr Pro Ala Val Gln Glu Leu Lys Ala Phe Ala Gly Glu Asp Phe Ser	
745 750 755	
CAA ATT CCT CGT GGT GAG AAA GAT TTA AAA CAA TTT GTA AGG CTA GGG	2359
Gln Ile Pro Arg Gly Glu Lys Asp Leu Lys Gln Phe Val Arg Leu Gly	
760 765 770	
CTA GAA ACA CAA AAT CAG ATT CAA CAA TTG GTT CAA CAA GGC TAT GAA	2407
Leu Glu Thr Gln Asn Gln Ile Gln Gln Leu Val Gln Gln Gly Tyr Glu	
775 780 785	
TTA GGG AAC GAA ATT GAC CTA TTA GAT GCT CAA GGA GAG TCC ACT GGT	2455
Leu Gly Asn Glu Ile Asp Leu Leu Asp Ala Gln Gly Glu Ser Thr Gly	
790 795 800	
GAA AAA GTT CGA CTT GAT ATT TCT AGT TTT GAG GAG TTT GGG GCT TTG	2503
Glu Lys Val Arg Leu Asp Ile Ser Ser Phe Glu Glu Phe Gly Ala Leu	
805 810 815 820	
CAA GCA AGG CTT AAT GAT GCT AGA GAA AAA TTG CCC GAA GAG GAG CGT	2551
Gln Ala Arg Leu Asn Asp Ala Arg Glu Lys Leu Pro Glu Glu Glu Arg	
825 830 835	

CCA GAC CTC ACA GAA ATT AGA GTG GGC TTA TCT TTA TAT GAC CAT GAG 2599
 Pro Asp Leu Thr Glu Ile Arg Val Gly Leu Ser Leu Tyr Asp His Glu
 840 845 850

ATT ATT GAC TAC GAT TGG CTG GTT GAT CTC TTA AAC CTC TTC ATG GAT 2647
 Ile Ile Asp Tyr Asp Trp Leu Val Asp Leu Leu Asn Leu Phe Met Asp
 855 860 865

CAG AAA ACT CCT GAA AAC AAA GCT TCA ATT GAG AAA CAT ATT CTC CCA 2695
 Gln Lys Thr Pro Glu Asn Lys Ala Ser Ile Glu Lys His Ile Leu Pro
 870 875 880

TTG GAT GAA ATG AGC CAA CAA GAG ATC AAG GAT ATT ATC GTA GAT ATA 2743
 Leu Asp Glu Met Ser Gln Gln Glu Ile Lys Asp Ile Ile Val Asp Ile
 885 890 895 900

GAA TCT GGG GAA ATT AAG GAA CAC TTT ACG AAA GCA ACC TTG GAA GAT 2791
 Glu Ser Gly Glu Ile Lys Glu His Phe Thr Lys Ala Thr Leu Glu Asp
 905 910 915

CAA AGA AAG CAC AAA CGC TCG GAT CGA CAA GAG TTA AAA ATC CGT CGA 2839
 Gln Arg Lys His Lys Arg Ser Asp Arg Gln Glu Leu Lys Ile Arg Arg
 920 925 930

TGG GCG GCC AAT CAA AAA GTC AAT GGC AAT CGT ATT GTC CAA GCC TTC 2887
 Trp Ala Ala Asn Gln Lys Val Asn Gly Asn Arg Ile Val Gln Ala Phe
 935 940 945

GAT CTT TTC TTA CCA GGT CAC AGC TTA GTA GAC AAT TCT CAA TTA TCA 2935
 Asp Leu Phe Leu Pro Gly His Ser Leu Val Asp Asn Ser Gln Leu Ser
 950 955 960

GCG CTT GTT TCA GAA ATT GAA GCA GAG GAA AAC TTA AGC TTT TTT GGA 2983
 Ala Leu Val Ser Glu Ile Glu Ala Glu Glu Asn Leu Ser Phe Phe Gly
 965 970 975 980

GCC TCA GAA TTT GAG ACT GCA TTA ATG AGC TTC TTC AAT TCA CTA TAA 3031
 Ala Ser Glu Phe Glu Thr Ala Leu Met Ser Phe Phe Asn Ser Leu *
 985 990 995

AAAAATAAAA ATACTCGCA 3050
 (2) INFORMATIONS POUR LA SEQ ID NO: 2:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 996 acides aminés
 (B) TYPE: acide aminé
 (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: protéine

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

Met Ala Glu Ala Lys Phe Glu Ala Ala Leu Ile Lys Lys Leu Glu Ala
 1 5 10 15

Glu Gly Trp Thr Tyr Arg Lys Asp Leu Ser Tyr Val Ser Ile Lys Val
 20 25 30

BNSDOCID: <FR 2767831A1 | >

Phe	Ala	Tyr	Ile	Gln	Phe	Lys	Asn	Arg	Ile	Lys	Lys	Val	Asn	Gly	Arg	
			340					345					350			
Glu	Leu	Lys	Arg	Glu	Leu	Ser	His	Lys	Gly	Ala	Ser	Gln	Ile	Leu	Leu	
		355					360					365				
Ile	Ser	Val	Gln	Gly	Leu	Thr	Lys	Ala	Val	Lys	Asn	Gly	Leu	Lys	Asn	
	370					375					380					
Thr	Asp	Arg	Asn	Val	Ile	Ile	Met	Asp	Glu	Ala	His	Arg	Ser	Ala	Asn	
385					390					395					400	
Gly	Glu	Ser	Val	Gln	Leu	Ile	Lys	Ser	Ala	Phe	Gln	Lys	Thr	Thr	Trp	
				405					410					415		
Phe	Gly	Phe	Thr	Gly	Thr	Pro	Asn	Phe	Tyr	Ser	Asp	Glu	Ile	Asn	Asp	
			420					425					430			
Val	Gln	Thr	Thr	Arg	Asp	Ile	Ser	Thr	His	Asp	Ile	Phe	Gly	Lys	Arg	
		435					440					445				
Leu	His	Ser	Tyr	Thr	Ile	Lys	Asp	Ala	Ile	Gly	Asp	Gly	Asn	Val	Leu	
	450					455					460					
Gly	Phe	Asp	Ile	Thr	Tyr	Phe	Asn	Pro	Thr	Ile	Glu	Ile	Glu	Ser	Leu	
465					470					475					480	
Asp	Glu	Glu	His	Ser	Glu	Lys	Asp	Tyr	Glu	Lys	Glu	Val	Tyr	Gln	Ser	
				485					490					495		
His	Val	Tyr	Arg	Glu	Gln	Val	Val	Gln	Asp	Ile	Leu	Asn	Leu	Trp	Asp	
			500					505					510			
Lys	Thr	Ser	Ser	Gly	Ala	Leu	Val	Ala	Gly	Lys	Arg	Glu	Lys	Asn	Val	
		515					520					525				
Phe	Gln	Ala	Met	Leu	Ala	Val	Ser	Gly	Lys	Gln	Ala	Val	Val	His	Tyr	
	530					535					540					
Tyr	Asn	Leu	Phe	Lys	Glu	Lys	Ala	Pro	His	Leu	Arg	Val	Ala	Met	Thr	
545					550					555					560	
Phe	Ser	Arg	Asp	Glu	Ser	Asn	Gln	Pro	Gly	Thr	Lys	Glu	Gln	Asn	Glu	
				565					570				575			
Ala	Leu	Lys	Lys	Ala	Ile	Lys	Glu	Tyr	Ser	Ser	Cys	Phe	Asn	Val	Pro	
			580					585					590			
Ser	Leu	Leu	Asn	Ala	Gln	Glu	Pro	Ala	Arg	Ala	Tyr	Met	Ile	Asp	Ile	
		595					600					605				
Thr	Lys	Arg	Leu	Ala	Arg	Lys	Lys	Pro	Tyr	Asn	Gln	Gly	Lys	Asp	Glu	
	610					615					620					
Asp	Arg	Leu	Asp	Leu	Val	Ile	Val	Ser	Asp	Gln	Leu	Leu	Thr	Gly	Phe	
625					630					635					640	

Asp Ser Lys Tyr Ile Asn Thr Ile Tyr Met Asp Lys Gln Leu Arg Glu
 645 650 655
 Gly Met Leu Ile Gln Ala Met Ser Arg Thr Asn Arg Thr Phe His Leu
 660 665 670
 Asn Ser Lys Pro His Gly Lys Val Arg Phe Tyr Arg Gln Gly Glu Gln
 675 680 685
 Met Lys Ser Phe Val Glu Asn Ala Leu Arg Ile Tyr Thr Arg Gly Gly
 690 695 700
 Asn Asp Thr Leu Gln Gly Ala Asp Glu Asp Ser Lys Asn Glu Asp Ile
 705 710 715 720
 Gln Ser Leu Glu Asp Glu His Ile Leu Ala Glu Pro Gln Ser His Gln
 725 730 735
 Ile Pro Lys Leu Thr Pro Ala Val Gln Glu Leu Lys Ala Phe Ala Gly
 740 745 750
 Glu Asp Phe Ser Gln Ile Pro Arg Gly Glu Lys Asp Leu Lys Gln Phe
 755 760 765
 Val Arg Leu Gly Leu Glu Thr Gln Asn Gln Ile Gln Gln Leu Val Gln
 770 775 780
 Gln Gly Tyr Glu Leu Gly Asn Glu Ile Asp Leu Leu Asp Ala Gln Gly
 785 790 795 800
 Glu Ser Thr Gly Glu Lys Val Arg Leu Asp Ile Ser Ser Phe Glu Glu
 805 810 815
 Phe Gly Ala Leu Gln Ala Arg Leu Asn Asp Ala Arg Glu Lys Leu Pro
 820 825 830
 Glu Glu Glu Arg Pro Asp Leu Thr Glu Ile Arg Val Gly Leu Ser Leu
 835 840 845
 Tyr Asp His Glu Ile Ile Asp Tyr Asp Trp Leu Val Asp Leu Leu Asn
 850 855 860
 Leu Phe Met Asp Gln Lys Thr Pro Glu Asn Lys Ala Ser Ile Glu Lys
 865 870 875 880
 His Ile Leu Pro Leu Asp Glu Met Ser Gln Gln Glu Ile Lys Asp Ile
 885 890 895
 Ile Val Asp Ile Glu Ser Gly Glu Ile Lys Glu His Phe Thr Lys Ala
 900 905 910
 Thr Leu Glu Asp Gln Arg Lys His Lys Arg Ser Asp Arg Gln Glu Leu
 915 920 925
 Lys Ile Arg Arg Trp Ala Ala Asn Gln Lys Val Asn Gly Asn Arg Ile
 930 935 940

Val Gln Ala Phe Asp Leu Phe Leu Pro Gly His Ser Leu Val Asp Asn
 945 950 955 960

Ser Gln Leu Ser Ala Leu Val Ser Glu Ile Glu Ala Glu Glu Asn Leu
 965 970 975

Ser Phe Phe Gly Ala Ser Glu Phe Glu Thr Ala Leu Met Ser Phe Phe
 980 985 990

Asn Ser Leu *
 995

(2) INFORMATIONS POUR LA SEQ ID NO: 3:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 3150 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: double
- (D) CONFIGURATION: linéaire

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: CDS
- (B) EMBLACEMENT: 39..3116

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 3:

TAAAAAATGG TGGCAACTAT GGAAATGAGG GGGATCAA ATG AGT CAT TCG GAA	53
Met Ser His Ser Glu	
1 5	
CAA ATG ATT GAA AAC CAG TTC ATA CAG ATC TTA AGT GAG AAA GAA AAT	101
Gln Met Ile Glu Asn Gln Phe Ile Gln Ile Leu Ser Glu Lys Glu Asn	
10 15 20	
CAG TGG ACT TAT CGT CCG GAC TTG AAG TCG GAA GAA GCA CTT TGC CAA	149
Gln Trp Thr Tyr Arg Pro Asp Leu Lys Ser Glu Glu Ala Leu Cys Gln	
25 30 35	
AAC TTT AGA GGC CAT TTG AAC CGA ATA AAT TTA GCA GTA TTG GAA GAG	197
Asn Phe Arg Gly His Leu Asn Arg Ile Asn Leu Ala Val Leu Glu Glu	
40 45 50	
CAA CTA TTA ACG GAC AAA GAA TTT AAG CAA GTC AAA GTC GAG TTT TCA	245
Gln-Leu-Leu-Thr-Asp-Lys-Glu-Phe-Lys-Gln-Val-Lys-Val-Glu-Phe-Ser	
55 60 65	
CGT TTG ACC GGA ACA CCT TTT TTA GCT TCT CAA TGG CTT AGA GGA GAA	293
Arg Leu Thr Gly Thr Pro Phe Leu Ala Ser Gln Trp Leu Arg Gly Glu	
70 75 80 85	
AAC GGG GTG GCT CAA GTT TTA TTA GAG CGA GAA GAT GGG GAA AAA GTG	341
Asn Gly Val Ala Gln Val Leu Leu Glu Arg Glu Asp Gly Glu Lys Val	
90 95 100	
ACT TTA GAA GCC TTT AGA AAT AAG GAT ATC TCA GGA GGA ACT TCT TCT	389
Thr Leu Glu Ala Phe Arg Asn Lys Asp Ile Ser Gly Gly Thr Ser Ser	
105 110 115	

TAT GAG GTA GTT CAC CAA GTG GTC CCA GAT TCC TCT AGA GTA GAT CGT Tyr Glu Val Val His Gln Val Val Pro Asp Ser Ser Arg Val Asp Arg 120 125 130	437
GGA GAT GTG AGC TTG CTG ATT AAT GGG CTC CCA ATC ATT CAT ATT GAG Gly Asp Val Ser Leu Leu Ile Asn Gly Leu Pro Ile Ile His Ile Glu 135 140 145	485
CTC CAA AAA GAG TCC GCT AAA GAC GGT TTC ATG CAA GCT TAT TAT CAA Leu Gln Lys Glu Ser Ala Lys Asp Gly Phe Met Gln Ala Tyr Tyr Gln 150 155 160 165	533
ATT CAG CGT TAT GCA GAA GAT GGA TTT TTT AAA GGG ATT TAC GCA ACC Ile Gln Arg Tyr Ala Glu Asp Gly Phe Phe Lys Gly Ile Tyr Ala Thr 170 175 180	581
ACT CAA ATC ATG GTG ATT CCC AAT AAA GTC GAT ACC CGA TAC TTT GCA Thr Gln Ile Met Val Ile Pro Asn Lys Val Asp Thr Arg Tyr Phe Ala 185 190 195	629
AGA CCT AGT GAA GAT ACC GCT GAA GCC TAT GCT CGG ATG AAG AAG TTT Arg Pro Ser Glu Asp Thr Ala Glu Ala Tyr Ala Arg Met Lys Lys Phe 200 205 210	677
TTA TTT AAT TGG CGG ACT GAA GAC AAT CAA ACG GTT TCC GAT TTG TTT Leu Phe Asn Trp Arg Thr Glu Asp Asn Gln Thr Val Ser Asp Leu Phe 215 220 225	725
GAT TTT ACT CGT ACA GTT TTG CGG ATA CCC GAT GCC CAT GAA TTG ATT Asp Phe Thr Arg Thr Val Leu Arg Ile Pro Asp Ala His Glu Leu Ile 230 235 240 245	773
AGC CAA TAT ACC ATT CTC GTC GAT GAT CCA AAA AAT CCA AAA TTC CTC Ser Gln Tyr Thr Ile Leu Val Asp Asp Pro Lys Asn Pro Lys Phe Leu 250 255 260	821
ATG GCT TTA AGG CCT TAC CAA ATT CAT GCT ATT CGT AAG ATT CGT CCA Met Ala Leu Arg Pro Tyr Gln Ile His Ala Ile Arg Lys Ile Arg Pro 265 270 275	869
AAA GCG GCA CAG CAT GAA GGA GGA TTC ATT TGG CAT GCG ACA GGT TCA Lys Ala Ala Gln His Glu Gly Gly Phe Ile Trp His Ala Thr Gly Ser 280 285 290	917
GGA AAG ACC ATT ACC AGT TTT GTC GCA ACG AAA TTA TTA GCA CAA AAT Gly Lys Thr Ile Thr Ser Phe Val Ala Thr Lys Leu Leu Ala Gln Asn 295 300 305	965
GCG ATC GGT GTC GAT CGT ACG GTC ATG GTT GTT GAT AGA ACA GAC TTA Ala Ile Gly Val Asp Arg Thr Val Met Val Val Asp Arg Thr Asp Leu 310 315 320 325	1013
GAT GCT CAA ACG CAG GAT GAG TTT ACG AAG TTT GCC TCG GAA TAC CAT Asp Ala Gln Thr Gln Asp Glu Phe Thr Lys Phe Ala Ser Glu Tyr His 330 335 340	1061

ACC GGA. CAA ACG ACC GGA AAT TCG GTA GCC AAT ACT TTG ATT GTT GGG	1109
Thr Gly Gln Thr Thr Gly Asn Ser Val Ala Asn Thr Leu Ile Val Gly	
345 350 355	
ATC AAA AAT CAA AAA CAG TTG GCT CGA AAC CTC CTA TCA TCA AAA AAT	1157
Ile Lys Asn Gln Lys Gln Leu Ala Arg Asn Leu Leu Ser Ser Lys Asn	
360 365 370	
AAT AAT ACG ATT TTA GTG ACC ACG ATT CAA AAA CTC TCT GCG GCT ATG	1205
Asn Asn Thr Ile Leu Val Thr Thr Ile Gln Lys Leu Ser Ala Ala Met	
375 380 385	
CGA AGT GCC CAA CAA GAG AGT GAA GAA AAA GGC TCG AAT CAA TTT GAG	1253
Arg Ser Ala Gln Gln Glu Ser Glu Glu Lys Gly Ser Asn Gln Phe Glu	
390 395 400 405	
AAG CTG CGG CAA GAA CAT ATT GTT TTT ATT GTT GAT GAG GCT CAT CGT	1301
Lys Leu Arg Gln Glu His Ile Val Phe Ile Val Asp Glu Ala His Arg	
410 415 420	
GCG GTT AGT GAT GAG GAA ATG AAA CGA ATT AAG AAA ATA TTA CCC AAT	1349
Ala Val Ser Asp Glu Glu Met Lys Arg Ile Lys Lys Ile Leu Pro Asn	
425 430 435	
TCT ACC TGG TTT GGA TTA ACG GGG ACG CCT ATT TTT GAA GCA AAT AAA	1397
Ser Thr Trp Phe Gly Leu Thr Gly Thr Pro Ile Phe Glu Ala Asn Lys	
440 445 450	
AAG CAA GAA AAT GGA ACC TTT GCC AGA ACG ACG AGC CAG CAA TAC GGG	1445
Lys Gln Glu Asn Gly Thr Phe Ala Arg Thr Thr Ser Gln Gln Tyr Gly	
455 460 465	
CCA CTC CTT CAC TCC TAT ACA ACC AAA AAT GCC ATG GAT GAT GGG GCT	1493
Pro Leu Leu His Ser Tyr Thr Thr Lys Asn Ala Met Asp Asp Gly Ala	
470 475 480 485	
GTG TTA GGC TTT CAA GTC GAG TAT CAT TCG CTG ATT TCA GAA GAG GAT	1541
Val Leu Gly Phe Gln Val Glu Tyr His Ser Leu Ile Ser Glu Glu Asp	
490 495 500	
CTA GAG GTG ATT GTC ACC CAA CTC AAT AAA GGA AAG TTG CCA GGC GAT	1589
Leu Glu Val Ile Val Thr Gln Leu Asn Lys Gly Lys Leu Pro Gly Asp	
505 510 515	
GCC CTC CAA CAA GAG GAA TTA TTG CCT GCG GAA CTT TAT GAA AAA GAT	1637
Ala Leu Gln Gln Glu Glu Leu Leu Pro Ala Glu Leu Tyr Glu Lys Asp	
520 525 530	
GAA CAT ATT CGT ACC ATG TTA CAA AAA ATC TTT AAT CGG AGG AGT GTG	1685
Glu His Ile Arg Thr Met Leu Gln Lys Ile Phe Asn Arg Arg Ser Val	
535 540 545	
GTC AAA AAG TTC AAG GTA AAG AAT GGT TTT CCT ACG ATG TCA GCC ATT	1733
Val Lys Lys Phe Lys Val Lys Asn Gly Phe Pro Thr Met Ser Ala Ile	
550 555 560 565	

CTT ACC ACT CAC TCG ATT GCT CAA GCC AAA CAT ATT TAC CGG ATT TTA	1781
Leu Thr Thr His Ser Ile Ala Gln Ala Lys His Ile Tyr Arg Ile Leu	
570 575 580	
AAG GAA ATG AAA GAT AAT GGG ACA CTC TTG AAT GGT CGA CAA TTT GAT	1829
Lys Glu Met Lys Asp Asn Gly Thr Leu Leu Asn Gly Arg Gln Phe Asp	
585 590 595	
GAA CGT CAT CGA CTG ATT GAT AAA GAT TTC CCA AGA GTA GCT ATT ACC	1877
Glu Arg His Arg Leu Ile Asp Lys Asp Phe Pro Arg Val Ala Ile Thr	
600 605 610	
TTC TCA ACC AAT CCG GAT CGA TTA GAA AAA AAT GAA CAA GAC GAT GAA	1925
Phe Ser Thr Asn Pro Asp Arg Leu Glu Lys Asn Glu Gln Asp Asp Glu	
615 620 625	
TTA GTA GAG ATC ATG AAA GAG TAT GCG AAA CAG TTT GAT GCT TCT CCT	1973
Leu Val Glu Ile Met Lys Glu Tyr Ala Lys Gln Phe Asp Ala Ser Pro	
630 635 640 645	
TAT CAA GAT GAG AAA CTC TAC AAT CAA AAT ATT AAT AAA CGG TTG GCC	2021
Tyr Gln Asp Glu Lys Leu Tyr Asn Gln Asn Ile Asn Lys Arg Leu Ala	
650 655 660	
CGT AAG GAA AAG CAA TAT CAA TCG GAT GGT CAG TGG CTA GAT TTT GTT	2069
Arg Lys Glu Lys Gln Tyr Gln Ser Asp Gly Gln Trp Leu Asp Phe Val	
665 670 675	
ATT GTT GTC GAT CGT TTA TTG ACA GGC TTT GAT TCT CCA GCG ATT CTA	2117
Ile Val Val Asp Arg Leu Leu Thr Gly Phe Asp Ser Pro Ala Ile Leu	
680 685 690	
ACC CTA TAT ATT GAT CGA GAA ATG AAC TAT CAA AAG CTC CTT CAA GCG	2165
Thr Leu Tyr Ile Asp Arg Glu Met Asn Tyr Gln Lys Leu Leu Gln Ala	
695 700 705	
TTC TCA AGA ACC AAC CGT ATT TAT ACA GGA AAA GAT TCT GGT TTG ATT	2213
Phe Ser Arg Thr Asn Arg Ile Tyr Thr Gly Lys Asp Ser Gly Leu Ile	
710 715 720 725	
GTC TCT TTT AGA AAA CCC TTC ACC ATG AAA GAG AAT GTG CAA AAC ACG	2261
Val Ser Phe Arg Lys Pro Phe Thr Met Lys Glu Asn Val Gln Asn Thr	
730 735 740	
TTT CGT CTA TTC TCA AAT GAA AAT CAA AAC TTT GAT CAA CTG ATT CCA	2309
Phe Arg Leu Phe Ser Asn Glu Asn Gln Asn Phe Asp Gln Leu Ile Pro	
745 750 755	
AGA GAA TAT GAA GAA GTC AAA AAA GAA TTT ATC GAA TGT TCA ACA CTT	2357
Arg Glu Tyr Glu Glu Val Lys Lys Glu Phe Ile Glu Cys Ser Thr Leu	
760 765 770	
TAT AAA CAA AGC GAA GCT GAC CTA TCG GAT AAT CCC CAC GAT CTT AAA	2405
Tyr Lys Gln Ser Glu Ala Asp Leu Ser Asp Asn Pro His Asp Leu Lys	
775 780 785	

ACG ATG ATT GCG CAA GTG AGT GCT TAT CAG AAG CTA GGT AAG AGT TAT	2453
Thr Met Ile Ala Gln Val Ser Ala Tyr Gln Lys Leu Gly Lys Ser Tyr	
790 795 800 805	
AAA GCG TTC CGA AGC TAT GAT CAA TAC GAA GAG GAC TTT GAA GCA TTT	2501
Lys Ala Phe Arg Ser Tyr Asp Gln Tyr Glu Glu Asp Phe Glu Ala Phe	
810 815 820	
TCA GAA GTG GTG GAG CAG TTG CCA CAA TAT CGA GGT AAA ACG GAA AAT	2549
Ser Glu Val Val Glu Gln Leu Pro Gln Tyr Arg Gly Lys Thr Glu Asn	
825 830 835	
GTT AAA ACA AAG ATT AAA GAA ATG ATC GAG GAT GAG GAG CAT CCT GAG	2597
Val Lys Thr Lys Ile Lys Glu Met Ile Glu Asp Glu Glu His Pro Glu	
840 845 850	
GAG GAC TTC GAG AAG CTT TTA CAA GAG ATT GCT TTT TCT TCG CAG CTC	2645
Glu Asp Phe Glu Lys Leu Leu Gln Glu Ile Ala Phe Ser Ser Gln Leu	
855 860 865	
AAT GCG ACA CAT AAA GAC GTC GTG GAT AGT TTT TAT ATT AAT CAA CTT	2693
Asn Ala Thr His Lys Asp Val Val Asp Ser Phe Tyr Ile Asn Gln Leu	
870 875 880 885	
TTG AAA GCG ATC CAA TTA AAC GAA GCA GGG GCC GTT GAA AAA TTT GAA	2741
Leu Lys Ala Ile Gln Leu Asn Glu Ala Gly Ala Val Glu Lys Phe Glu	
890 895 900	
AAA GAA ATC CAA CAA AAG GAT CCT CAA ATT CAA AAG ATG TAT CAC ACC	2789
Lys Glu Ile Gln Gln Lys Asp Pro Gln Ile Gln Lys Met Tyr His Thr	
905 910 915	
TTG AAA GAT CAA CTC GTC AAT ACT ACT GAA GAG ATC GAT GTG GCT CAA	2837
Leu Lys Asp Gln Leu Val Asn Thr Thr Glu Glu Ile Asp Val Ala Gln	
920 925 930	
TTA AAA GAG ACT TCG ATT CAA AAT GAA ATT CAA AGA CTA CTT CAA AAA	2885
Leu Lys Glu Thr Ser Ile Gln Asn Glu Ile Gln Arg Leu Leu Gln Lys	
935 940 945	
GAA GCT GAA GAA TTT GGA TTA TCC TTC GAC TTT TTA CAG TCT GCA ATG	2933
Glu Ala Glu Glu Phe Gly Leu Ser Phe Asp Phe Leu Gln Ser Ala Met	
950 955 960 965	
AAT GAA TAT CAA GGC GAT AAG AAA GCG ATC CCT TAT TTA ACC CAT TTA	2981
Asn Glu Tyr Gln Gly Asp Lys Lys Ala Ile Pro Tyr Leu Thr His Leu	
970 975 980	
CTG GAT TCG ATG ACT CTT AGT AAA GAA GAG TTC GAA CCC AAA GCG GGT	3029
Leu Asp Ser Met Thr Leu Ser Lys Glu Glu Phe Glu Pro Lys Ala Gly	
985 990 995	
GAA AAA TAC AGA AGA AGA CCA AAA GTT TTA GAA GAA CGA CTA CGA CAA	3077
Glu Lys Tyr Arg Arg Arg Pro Lys Val Leu Glu Glu Arg Leu Arg Gln	
1000 1005 1010	

AAC TTT GAA CAA ATT CAA AAA TGG AAA GAA GAA TTA TAA TGGCGACAGG 3126
 Asn Phe Glu Gln Ile Gln Lys Trp Lys Glu Glu Leu *
 1015 1020 1025

TTTAAATCAA CAACTATGGG CTTC 3150

(2) INFORMATIONS POUR LA SEQ ID NO: 4:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 1026 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: protéine

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 4:

Met Ser His Ser Glu Gln Met Ile Glu Asn Gln Phe Ile Gln Ile Leu
 1 5 10 15
 Ser Glu Lys Glu Asn Gln Trp Thr Tyr Arg Pro Asp Leu Lys Ser Glu
 20 25 30
 Glu Ala Leu Cys Gln Asn Phe Arg Gly His Leu Asn Arg Ile Asn Leu
 35 40 45
 Ala Val Leu Glu Glu Gln Leu Leu Thr Asp Lys Glu Phe Lys Gln Val
 50 55 60
 Lys Val Glu Phe Ser Arg Leu Thr Gly Thr Pro Phe Leu Ala Ser Gln
 65 70 75 80
 Trp Leu Arg Gly Glu Asn Gly Val Ala Gln Val Leu Leu Glu Arg Glu
 85 90 95
 Asp Gly Glu Lys Val Thr Leu Glu Ala Phe Arg Asn Lys Asp Ile Ser
 100 105 110
 Gly Gly Thr Ser Ser Tyr Glu Val Val His Gln Val Val Pro Asp Ser
 115 120 125
 Ser Arg Val Asp Arg Gly Asp Val Ser Leu Leu Ile Asn Gly Leu Pro
 130 135 140
 Ile Ile His Ile Glu Leu Gln Lys Glu Ser Ala Lys Asp Gly Phe Met
 145 150 155 160
 Gln Ala Tyr Tyr Gln Ile Gln Arg Tyr Ala Glu Asp Gly Phe Phe Lys
 165 170 175
 Gly Ile Tyr Ala Thr Thr Gln Ile Met Val Ile Pro Asn Lys Val Asp
 180 185 190
 Thr Arg Tyr Phe Ala Arg Pro Ser Glu Asp Thr Ala Glu Ala Tyr Ala
 195 200 205
 Arg Met Lys Lys Phe Leu Phe Asn Trp Arg Thr Glu Asp Asn Gln Thr
 210 215 220

Val	Ser	Asp	Leu	Phe	Asp	Phe	Thr	Arg	Thr	Val	Leu	Arg	Ile	Pro	Asp	225	230	235	240
Ala	His	Glu	Leu	Ile	Ser	Gln	Tyr	Thr	Ile	Leu	Val	Asp	Asp	Pro	Lys	245	250	255	
Asn	Pro	Lys	Phe	Leu	Met	Ala	Leu	Arg	Pro	Tyr	Gln	Ile	His	Ala	Ile	260	265	270	
Arg	Lys	Ile	Arg	Pro	Lys	Ala	Ala	Gln	His	Glu	Gly	Gly	Phe	Ile	Trp	275	280	285	
His	Ala	Thr	Gly	Ser	Gly	Lys	Thr	Ile	Thr	Ser	Phe	Val	Ala	Thr	Lys	290	295	300	
Leu	Leu	Ala	Gln	Asn	Ala	Ile	Gly	Val	Asp	Arg	Thr	Val	Met	Val	Val	305	310	315	320
Asp	Arg	Thr	Asp	Leu	Asp	Ala	Gln	Thr	Gln	Asp	Glu	Phe	Thr	Lys	Phe	325	330	335	
Ala	Ser	Glu	Tyr	His	Thr	Gly	Gln	Thr	Thr	Gly	Asn	Ser	Val	Ala	Asn	340	345	350	
Thr	Leu	Ile	Val	Gly	Ile	Lys	Asn	Gln	Lys	Gln	Leu	Ala	Arg	Asn	Leu	355	360	365	
Leu	Ser	Ser	Lys	Asn	Asn	Asn	Thr	Ile	Leu	Val	Thr	Thr	Ile	Gln	Lys	370	375	380	
Leu	Ser	Ala	Ala	Met	Arg	Ser	Ala	Gln	Gln	Glu	Ser	Glu	Glu	Lys	Gly	385	390	395	400
Ser	Asn	Gln	Phe	Glu	Lys	Leu	Arg	Gln	Glu	His	Ile	Val	Phe	Ile	Val	405	410	415	
Asp	Glu	Ala	His	Arg	Ala	Val	Ser	Asp	Glu	Glu	Met	Lys	Arg	Ile	Lys	420	425	430	
Lys	Ile	Leu	Pro	Asn	Ser	Thr	Trp	Phe	Gly	Leu	Thr	Gly	Thr	Pro	Ile	435	440	445	
Phe	Glu	Ala	Asn	Lys	Lys	Gln	Glu	Asn	Gly	Thr	Phe	Ala	Arg	Thr	Thr	450	455	460	
Ser	Gln	Gln	Tyr	Gly	Pro	Leu	Leu	His	Ser	Tyr	Thr	Thr	Lys	Asn	Ala	465	470	475	480
Met	Asp	Asp	Gly	Ala	Val	Leu	Gly	Phe	Gln	Val	Glu	Tyr	His	Ser	Leu	485	490	495	
Ile	Ser	Glu	Glu	Asp	Leu	Glu	Val	Ile	Val	Thr	Gln	Leu	Asn	Lys	Gly	500	505	510	
Lys	Leu	Pro	Gly	Asp	Ala	Leu	Gln	Gln	Glu	Glu	Leu	Leu	Pro	Ala	Glu	515	520	525	

Leu Tyr Glu Lys Asp Glu His Ile Arg Thr Met Leu Gln Lys Ile Phe
 530 535 540
 Asn Arg Arg Ser Val Val Lys Lys Phe Lys Val Lys Asn Gly Phe Pro
 545 550 555 560
 Thr Met Ser Ala Ile Leu Thr Thr His Ser Ile Ala Gln Ala Lys His
 565 570 575
 Ile Tyr Arg Ile Leu Lys Glu Met Lys Asp Asn Gly Thr Leu Leu Asn
 580 585 590
 Gly Arg Gln Phe Asp Glu Arg His Arg Leu Ile Asp Lys Asp Phe Pro
 595 600 605
 Arg Val Ala Ile Thr Phe Ser Thr Asn Pro Asp Arg Leu Glu Lys Asn
 610 615 620
 Glu Gln Asp Asp Glu Leu Val Glu Ile Met Lys Glu Tyr Ala Lys Gln
 625 630 635 640
 Phe Asp Ala Ser Pro Tyr Gln Asp Glu Lys Leu Tyr Asn Gln Asn Ile
 645 650 655
 Asn Lys Arg Leu Ala Arg Lys Glu Lys Gln Tyr Gln Ser Asp Gly Gln
 660 665 670
 Trp Leu Asp Phe Val Ile Val Val Asp Arg Leu Leu Thr Gly Phe Asp
 675 680 685
 Ser Pro Ala Ile Leu Thr Leu Tyr Ile Asp Arg Glu Met Asn Tyr Gln
 690 695 700
 Lys Leu Leu Gln Ala Phe Ser Arg Thr Asn Arg Ile Tyr Thr Gly Lys
 705 710 715 720
 Asp Ser Gly Leu Ile Val Ser Phe Arg Lys Pro Phe Thr Met Lys Glu
 725 730 735
 Asn Val Gln Asn Thr Phe Arg Leu Phe Ser Asn Glu Asn Gln Asn Phe
 740 745 750
 Asp Gln Leu Ile Pro Arg Glu Tyr Glu Glu Val Lys Lys Glu Phe Ile
 755 760 765
 Glu Cys Ser Thr Leu Tyr Lys Gln Ser Glu Ala Asp Leu Ser Asp Asn
 770 775 780
 Pro His Asp Leu Lys Thr Met Ile Ala Gln Val Ser Ala Tyr Gln Lys
 785 790 795 800
 Leu Gly Lys Ser Tyr Lys Ala Phe Arg Ser Tyr Asp Gln Tyr Glu Glu
 805 810 815
 Asp Phe Glu Ala Phe Ser Glu Val Val Glu Gln Leu Pro Gln Tyr Arg
 820 825 830

Gly Lys Thr Glu Asn Val Lys Thr Lys Ile Lys Glu Met Ile Glu Asp
 835 840 845

 Glu Glu His Pro Glu Glu Asp Phe Glu Lys Leu Leu Gln Glu Ile Ala
 850 855 860

 Phe Ser Ser Gln Leu Asn Ala Thr His Lys Asp Val Val Asp Ser Phe
 865 870 875 880

 Tyr Ile Asn Gln Leu Leu Lys Ala Ile Gln Leu Asn Glu Ala Gly Ala
 885 890 895

 Val Glu Lys Phe Glu Lys Glu Ile Gln Gln Lys Asp Pro Gln Ile Gln
 900 905 910

 Lys Met Tyr His Thr Leu Lys Asp Gln Leu Val Asn Thr Thr Glu Glu
 915 920 925

 Ile Asp Val Ala Gln Leu Lys Glu Thr Ser Ile Gln Asn Glu Ile Gln
 930 935 940

 Arg Leu Leu Gln Lys Glu Ala Glu Glu Phe Gly Leu Ser Phe Asp Phe
 945 950 955 960

 Leu Gln Ser Ala Met Asn Glu Tyr Gln Gly Asp Lys Lys Ala Ile Pro
 965 970 975

 Tyr Leu Thr His Leu Leu Asp Ser Met Thr Leu Ser Lys Glu Glu Phe
 980 985 990

 Glu Pro Lys Ala Gly Glu Lys Tyr Arg Arg Arg Pro Lys Val Leu Glu
 995 1000 1005

 Glu Arg Leu Arg Gln Asn Phe Glu Gln Ile Gln Lys Trp Lys Glu Glu
 1010 1015 1020

 Leu *
 . 1025

(2) INFORMATIONS POUR LA SEQ ID NO: 5:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1600 paires de bases
 (B) TYPE: nucléotide
 (C) NOMBRE DE BRINS: double
 (D) CONFIGURATION: linéaire

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: CDS
 (B) EMBLACEMENT: 21..1568
 (D) AUTRES INFORMATIONS: /product= "HsdM"

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 5:

AGAAATAAGG AGCGAATAAAA ATG GCA TTA TCA AAT GAA CAA AAA AAT AAA
 Met Ala Leu Ser Asn Glu Gln Lys Asn Lys
 1 5 10

50

ATG TGG GCC CTC TTA AAT CAA ACC CGA GGA CAG ATT GGT TTA ACC GCC	98
Met Trp Ala Leu Leu Asn Gln Thr Arg Gly Gln Ile Gly Leu Thr Ala	
15 20 25	
TAT AAA GAC TAC ATT TTT GGC TTA CTA TTT TAT AAA TAT TTA TCA GAA	146
Tyr Lys Asp Tyr Ile Phe Gly Leu Leu Phe Tyr Lys Tyr Leu Ser Glu	
30 35 40	
AAG GCC ACA CAA TGG CTG GGA GAA GTT TTA CGA GGA GAC ACT TGG GAA	194
Lys Ala Thr Gln Trp Leu Gly Glu Val Leu Arg Gly Asp Thr Trp Glu	
45 50 55	
AAT GTT TAT GAT CAA GAT CCT GTA AGA GCC TTA GAC TAT ATG AAA CAA	242
Asn Val Tyr Asp Gln Asp Pro Val Arg Ala Leu Asp Tyr Met Lys Gln	
60 65 70	
AAG CTC GGT TAT GCC ATC CAA CCG AAG GAA TTT TTT AAA GAT TGG GAA	290
Lys Leu Gly Tyr Ala Ile Gln Pro Lys Glu Phe Phe Lys Asp Trp Glu	
75 80 85 90	
GCT GCC ATT CAT GAA GAA CGA TTT AAC ATT CCA ATG ATT TCA GAT ACA	338
Ala Ala Ile His Glu Glu Arg Phe Asn Ile Pro Met Ile Ser Asp Thr	
95 100 105	
TTT GGC CAT TTT AAT CAG CAA ATT GCT TTT GAA GCT AAA GAT GAT TTT	386
Phe Gly His Phe Asn Gln Gln Ile Ala Phe Glu Ala Lys Asp Asp Phe	
110 115 120	
GAA GGC ATC TTT GAT GGG ATG CGT TTT GAT AGT TCA GAT TTA GGA AGT	434
Glu Gly Ile Phe Asp Gly Met Arg Phe Asp Ser Ser Asp Leu Gly Ser	
125 130 135	
AAT GCG CAA GCT CGA GCA AGT GTG ATG ATC TCG ATG ATT GAG TTA CTC	482
Asn Ala Gln Ala Arg Ala Ser Val Met Ile Ser Met Ile Glu Leu Leu	
140 145 150	
TCT GCC CCT GAG TTT GAT CTC TCT ACA GGA GGG GAC ACC GTT TCA GAT	530
Ser Ala Pro Glu Phe Asp Leu Ser Thr Gly Gly Asp Thr Val Ser Asp	
155 160 165 170	
ATC TAC GAA TAC TTG TTA GAA AAG TTT GCG ACC GTA TTG GCC TCA GAC	578
Ile Tyr Glu Tyr Leu Leu Glu Lys Phe Ala Thr Val Leu Ala Ser Asp	
175 180 185	
ATG GGA CAA TAC TAT ACT CCC AAA GAA ATT TCA GAG GTC ATG GCT CGT	626
Met Gly Gln Tyr Tyr Thr Pro Lys Glu Ile Ser Glu Val Met Ala Arg	
190 195 200	
ATT TTG ACT TTT GGT AAA GCC GAC GAA GAC AAT TTT TCT ATT TAT GAT	674
Ile Leu Thr Phe Gly Lys Ala Asp Glu Asp Asn Phe Ser Ile Tyr Asp	
205 210 215	
CCC GCG GTC GGT TCA GCT TCT CTC TTA ATT ACT ACC GCC AGT CAT ATG	722
Pro Ala Val Gly Ser Ala Ser Leu Leu Ile Thr Thr Ala Ser His Met	
220 225 230	

AAA CAT TCC AAT CAA AGG GGA GCG ATT AAA TAC TTT GGT CAA GAA AAA	770
Lys His Ser Asn Gln Arg Gly Ala Ile Lys Tyr Phe Gly Gln Glu Lys	
235 240 245 250	
GAT GCC ACC CCT TAC CGA TTG GCT CGA ATG AAT TTA ATG ATG CAT AAC	818
Asp Ala Thr Pro Tyr Arg Leu Ala Arg Met Asn Leu Met Met His Asn	
255 260 265	
ATT GAA TAC AAT GAT ATT CAG ATC CAT CAT GCG GAT ACT TTG GAA AGC	866
Ile Glu Tyr Asn Asp Ile Gln Ile His His Ala Asp Thr Leu Glu Ser	
270 275 280	
GAT TGG CCT GAC GGA GTC ATT GAA GGG AAA GAT ACG CCT CGA ATG TTT	914
Asp Trp Pro Asp Gly Val Ile Glu Gly Lys Asp Thr Pro Arg Met Phe	
285 290 295	
GAC GCA GTG ATG GCT AAT CCA CCT TAT TCT GCC CAT TGG AAT AAT AAG	962
Asp Ala Val Met Ala Asn Pro Pro Tyr Ser Ala His Trp Asn Asn Lys	
300 305 310	
GAT CGG GAA GAT GAC CCT CGA TTT CGT GAA TAT GGC ATT GCG CCA AAA	1010
Asp Arg Glu Asp Asp Pro Arg Phe Arg Glu Tyr Gly Ile Ala Pro Lys	
315 320 325 330	
ACT AAA GCG GAC TAT TCT TTC TTG TTG CAT TGT TTG TAC CAT ACC AAA	1058
Thr Lys Ala Asp Tyr Ser Phe Leu Leu His Cys Leu Tyr His Thr Lys	
335 340 345	
GAG AGT GGC CGT GTG GCC ATT ATT CTC CCA CAC GGG GTC TTG TTT AGA	1106
Glu Ser Gly Arg Val Ala Ile Ile Leu Pro His Gly Val Leu Phe Arg	
350 355 360	
GGC GCT GCG GAA GGA CGG ATT CGA AAA GCT TTG ATT GAT AAA CAT CAA	1154
Gly Ala Ala Glu Gly Arg Ile Arg Lys Ala Leu Ile Asp Lys His Gln	
365 370 375	
ATT GAA GCG GTA ATT GGT TTT CCA GAT AAA CTC TTT TTG AAC ACA GGC	1202
Ile Glu Ala Val Ile Gly Phe Pro Asp Lys Leu Phe Leu Asn Thr Gly	
380 385 390	
ATT CCC GTT TGT GTC TTA ATT CTT AAA AAG AAT CGA GCA AAC TCT GAT	1250
Ile Pro Val Cys Val Leu Ile Leu Lys Lys Asn Arg Ala Asn Ser Asp	
395 400 405 410	
ATT CTT TTT GTG GAC GCC AGT CAA GGC TTT GAA AAG ATG AAG AAT CAA	1298
Ile Leu Phe Val Asp Ala Ser Gln Gly Phe Glu Lys Met Lys Asn Gln	
415 420 425	
AAA CAA TTG CGC CCA GAA GAT ATT GAC AAG ATC ACT GAA ACA GTC ATT	1346
Lys Gln Leu Arg Pro Glu Asp Ile Asp Lys Ile Thr Glu Thr Val Ile	
430 435 440	
CAT CGC AAG GCG GTT GAT AAA TAT AGC CAC TTA GCA ACG CTT GAG GAA	1394
His Arg Lys Ala Val Asp Lys Tyr Ser His Leu Ala Thr Leu Glu Glu	
445 450 455	

GTC ATT GAG AAT GAT TAC AAC TTA AAT ATT CCG CGT TAT GTG GAT ACC 1442
 Val Ile Glu Asn Asp Tyr Asn Leu Asn Ile Pro Arg Tyr Val Asp Thr
 460 465 470

TTT GAA GAG GAA GAA TCG ATT GAC TTA GCG GAT ATT CAA GGT CAA ATA 1490
 Phe Glu Glu Glu Glu Ser Ile Asp Leu Ala Asp Ile Gln Gly Gln Ile
 475 480 485 490

GAT GAA GTA GAC GCA GAG ATT GCG AAA GCT AAT CAA ACC TTA GCA AAC 1538
 Asp Glu Val Asp Ala Glu Ile Ala Lys Ala Asn Gln Thr Leu Ala Asn
 495 500 505

TAC TTT AAG GAA CTG GGA GTG CTA AAA TGA AGGAAAGATT AAAGGCTCCA 1588
 Tyr Phe Lys Glu Leu Gly Val Leu Lys *
 510 515

GAATTAAGGT TT 1600

(2) INFORMATIONS POUR LA SEQ ID NO: 6:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 516 acides aminés
- (B) TYPE: acide aminé
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: protéine

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 6:

Met Ala Leu Ser Asn Glu Gln Lys Asn Lys Met Trp Ala Leu Leu Asn
 1 5 10 15

Gln Thr Arg Gly Gln Ile Gly Leu Thr Ala Tyr Lys Asp Tyr Ile Phe
 20 25 30

Gly Leu Leu Phe Tyr Lys Tyr Leu Ser Glu Lys Ala Thr Gln Trp Leu
 35 40 45

Gly Glu Val Leu Arg Gly Asp Thr Trp Glu Asn Val Tyr Asp Gln Asp
 50 55 60

Pro Val Arg Ala Leu Asp Tyr Met Lys Gln Lys Leu Gly Tyr Ala Ile
 65 70 75 80

Gln Pro Lys Glu Phe Phe Lys Asp Trp Glu Ala Ala Ile His Glu Glu
 85 90 95

Arg Phe Asn Ile Pro Met Ile Ser Asp Thr Phe Gly His Phe Asn Gln
 100 105 110

Gln Ile Ala Phe Glu Ala Lys Asp Asp Phe Glu Gly Ile Phe Asp Gly
 115 120 125

Met Arg Phe Asp Ser Ser Asp Leu Gly Ser Asn Ala Gln Ala Arg Ala
 130 135 140

Ser Val Met Ile Ser Met Ile Glu Leu Leu Ser Ala Pro Glu Phe Asp
 145 150 155 160

Leu Ser Thr Gly Gly Asp Thr Val Ser Asp Ile Tyr Glu Tyr Leu Leu
 165 170 175
 Glu Lys Phe Ala Thr Val Leu Ala Ser Asp Met Gly Gln Tyr Tyr Thr
 180 185 190
 Pro Lys Glu Ile Ser Glu Val Met Ala Arg Ile Leu Thr Phe Gly Lys
 195 200 205
 Ala Asp Glu Asp Asn Phe Ser Ile Tyr Asp Pro Ala Val Gly Ser Ala
 210 215 220
 Ser Leu Leu Ile Thr Thr Ala Ser His Met Lys His Ser Asn Gln Arg
 225 230 235 240
 Gly Ala Ile Lys Tyr Phe Gly Gln Glu Lys Asp Ala Thr Pro Tyr Arg
 245 250 255
 Leu Ala Arg Met Asn Leu Met Met His Asn Ile Glu Tyr Asn Asp Ile
 260 265 270
 Gln Ile His His Ala Asp Thr Leu Glu Ser Asp Trp Pro Asp Gly Val
 275 280 285
 Ile Glu Gly Lys Asp Thr Pro Arg Met Phe Asp Ala Val Met Ala Asn
 290 295 300
 Pro Pro Tyr Ser Ala His Trp Asn Asn Lys Asp Arg Glu Asp Asp Pro
 305 310 315 320
 Arg Phe Arg Glu Tyr Gly Ile Ala Pro Lys Thr Lys Ala Asp Tyr Ser
 325 330 335
 Phe Leu Leu His Cys Leu Tyr His Thr Lys Glu Ser Gly Arg Val Ala
 340 345 350
 Ile Ile Leu Pro His Gly Val Leu Phe Arg Gly Ala Ala Glu Gly Arg
 355 360 365
 Ile Arg Lys Ala Leu Ile Asp Lys His Gln Ile Glu Ala Val Ile Gly
 370 375 380
 Phe Pro Asp Lys Leu Phe Leu Asn Thr Gly Ile Pro Val Cys Val Leu
 385 390 395 400
 Ile Leu Lys Lys Asn Arg Ala Asn Ser Asp Ile Leu Phe Val Asp Ala
 405 410 415
 Ser Gln Gly Phe Glu Lys Met Lys Asn Gln Lys Gln Leu Arg Pro Glu
 420 425 430
 Asp Ile Asp Lys Ile Thr Glu Thr Val Ile His Arg Lys Ala Val Asp
 435 440 445
 Lys Tyr Ser His Leu Ala Thr Leu Glu Glu Val Ile Glu Asn Asp Tyr
 450 455 460

Asn Leu Asn Ile Pro Arg Tyr Val Asp Thr Phe Glu Glu Glu Glu Ser
 465 470 475 480

Ile Asp Leu Ala Asp Ile Gln Gly Gln Ile Asp Glu Val Asp Ala Glu
 485 490 495

Ile Ala Lys Ala Asn Gln Thr Leu Ala Asn Tyr Phe Lys Glu Leu Gly
 500 505 510

Val Leu Lys *
 515

(2) INFORMATIONS POUR LA SEQ ID NO: 7:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1700 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: double
- (D) CONFIGURATION: linéaire

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: CDS
- (B) EMBLEMMENT: 48..1661
- (D) AUTRES INFORMATIONS: /product= "hdsM"

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 7:

AAAGTTTTAG AAGAACGACT ACGACAAAAC TTTGAACAAA TTCAAAA ATG GAA AGA 56
 Met Glu Arg
 1

AGA ATT ATA ATG GCG ACA GGT TTA AAT CAA CAA CTA TGG GCT TCT GCG 104
 Arg Ile Ile Met Ala Thr Gly Leu Asn Gln Gln Leu Trp Ala Ser Ala
 5 10 15

GAT ATC TTA CGT GGG AAG ATG GAT GCA AGT GAG TAC AAA AAC TAT TTA 152
 Asp Ile Leu Arg Gly Lys Met Asp Ala Ser Glu Tyr Lys Asn Tyr Leu
 20 25 30 35

CTG GGA TTA ATT TTC TAT AAG TAC TTG TCT GAT GCG CAG CTG AGA GAA 200
 Leu Gly Leu Ile Phe Tyr Lys Tyr Leu Ser Asp Ala Gln Leu Arg Glu
 40 45 50

GTT TAT GAA CAA GAA AAT GGG AAG ACA GAT ACC TTC CCA GAA CGT TCC 248
 Val Tyr Glu Gln Glu Asn Gly Lys Thr Asp Thr Phe Pro Glu Arg Ser
 55 60 65

ACT CAG TAT GCA GGC TTC ATG GAA TGG TAC GAA GAA GAC AAA GAC GAT 296
 Thr Gln Tyr Ala Gly Phe Met Glu Trp Tyr Glu Glu Asp Lys Asp Asp
 70 75 80

TTA ATC GAA AAT ATC CAA CCC AAA CAA GGT TAC TTT ATC CAA CCC GAT 344
 Leu Ile Glu Asn Ile Gln Pro Lys Gln Gly Tyr Phe Ile Gln Pro Asp
 85 90 95

CAA TTG TTT TAT AGT TAC CGC ATC AAA GCA GAC AAC TAT GAA TTT AAT 392
 Gln Leu Phe Tyr Ser Tyr Arg Ile Lys Ala Asp Asn Tyr Glu Phe Asn
 100 105 110 115

TTG ACC GAT TTA CAA GCA GGT TTT AAT GAG TTG GAA CGT CAA GGA GAA Leu Thr Asp Leu Gln Ala Gly Phe Asn Glu Leu Glu Arg Gln Gly Glu 120 125 130	440
GAA TTC AGT GGA TTG TTT GCG GAC ATT GAT TTA AAC TCC ACA AAA TTA Glu Phe Ser Gly Leu Phe Ala Asp Ile Asp Leu Asn Ser Thr Lys Leu 135 140 145	488
GGA TCA AAT GCA CTA CTA CGT AAT GTG ACG ATC ACT GAG GTT TTG CGT Gly Ser Asn Ala Leu Leu Arg Asn Val Thr Ile Thr Glu Val Leu Arg 150 155 160	536
GCC TTA GAT GAG ATT GAT TTA TTT GAA CAC AAT GGC GAT GTG ATT GGG Ala Leu Asp Glu Ile Asp Leu Phe Glu His Asn Gly Asp Val Ile Gly 165 170 175	584
GAC GCT TAT GAG TAT TTG ATT GGA GAA TTC GCG TCA AGT GCA GGT AAA Asp Ala Tyr Glu Tyr Leu Ile Gly Glu Phe Ala Ser Ser Ala Gly Lys 180 185 190 195	632
AAA GCC GGT GAG TTT TAT ACG CCT CAA GCT GTT TCG AAA ATC ATG TCA Lys Ala Gly Glu Phe Tyr Thr Pro Gln Ala Val Ser Lys Ile Met Ser 200 205 210	680
GAA ATT ACG TCG ATT GGA CAA GAA ACC CGT GCG CCT TTT CAT ATT TAT Glu Ile Thr Ser Ile Gly Gln Glu Thr Arg Ala Pro Phe His Ile Tyr 215 220 225	728
GAT CCT GCG ATG GGG TCA GGT TCT TTG ATG CTT AAT ATC CGT CGG TAT Asp Pro Ala Met Gly Ser Gly Ser Leu Met Leu Asn Ile Arg Arg Tyr 230 235 240	776
CTC AAT AAT CCA GAT CAA GTG CAC TAT CAT GGA CAA GAG TTA AAT ACG Leu Asn Asn Pro Asp Gln Val His Tyr His Gly Gln Glu Leu Asn Thr 245 250 255	824
ACG ACC TTT AAC TTA GCA CGT ATG AAC TTA ATT CTT CAT GGA ATA GAT Thr Thr Phe Asn Leu Ala Arg Met Asn Leu Ile Leu His Gly Ile Asp 260 265 270 275	872
AAA GAA CGC ATG AAC CTG AAT AAC GGA GAT ACT TTA GAT GCC GAT TGG Lys Glu Arg Met Asn Leu Asn Asn Gly Asp Thr Leu Asp Ala Asp Trp 280 285 290	920
CCC TCA GAA GAA CCG TAT CAG TTT GAT TCC GTA TGC ATG AAC CCT CCT Pro Ser Glu Glu Pro Tyr Gln Phe Asp Ser Val Cys Met Asn Pro Pro 295 300 305	968
TAT TCT GCG AAA TGG TCA GCG GCA GAT CAA TTT CTC TCT GAC CCT CGT Tyr Ser Ala Lys Trp Ser Ala Ala Asp Gln Phe Leu Ser Asp Pro Arg 310 315 320	1016
TTT GAG CGT TTT GGA AAA TTA GCG CCT AAA TCT AAA GCG GAC TTT GCC Phe Glu Arg Phe Gly Lys Leu Ala Pro Lys Ser Lys Ala Asp Phe Ala 325 330 335	1064

TTT CTC CTT CAC GGT TTT TAC CAT TTG AAA GAA TCG GGA ACA ATG GGG Phe Leu Leu His Gly Phe Tyr His Leu Lys Glu Ser Gly Thr Met Gly 340 345 350 355	1112
ATT GTC CTG CCT CAT GGC GTT CTC TTT AGA GGA GCC GCG GAA GGA ACC Ile Val Leu Pro His Gly Val Leu Phe Arg Gly Ala Ala Glu Gly Thr 360 365 370	1160
ATC CGT CAA GCC CTT TTA GAA ATG GGA GCT ATT GAT GCG GTC ATT GGC Ile Arg Gln Ala Leu Leu Glu Met Gly Ala Ile Asp Ala Val Ile Gly 375 380 385	1208
TTG CCG GCC AAT ATC TTT TTT GGA ACG AGT ATT CCG ACA ACG GTT ATT Leu Pro Ala Asn Ile Phe Phe Gly Thr Ser Ile Pro Thr Thr Val Ile 390 395 400	1256
ATT TTG AAG AGA AAC CGC TCT CGT CGT GAT GTC TTA TTT ATC GAT GCT Ile Leu Lys Arg Asn Arg Ser Arg Arg Asp Val Leu Phe Ile Asp Ala 405 410 415	1304
TCT CAA GAC TTT GAA AAA CGA AAA AAT CAA AAT GTC CTG TTG GAT GAA Ser Gln Asp Phe Glu Lys Arg Lys Asn Gln Asn Val Leu Leu Asp Glu 420 425 430 435	1352
CAT ATT GAT AAA ATT GTT TCT ATC CAC AAA AAA CGA GAA GAT ATT GAA His Ile Asp Lys Ile Val Ser Ile His Lys Lys Arg Glu Asp Ile Glu 440 445 450	1400
AGA TAT GCT CAT GTT GCA AGT TTT GAT GAG ATC CAA GAA AAT GAC TTT Arg Tyr Ala His Val Ala Ser Phe Asp Glu Ile Gln Glu Asn Asp Phe 455 460 465	1448
AAC TTA AAT ATC CCT CGT TAT GTA GAT ACC TTT GAG GAA GAG GAA CCG Asn Leu Asn Ile Pro Arg Tyr Val Asp Thr Phe Glu Glu Glu Glu Pro 470 475 480	1496
GTT GAT TTG GTT GCA GTA AAT ACC AAT CTC CTT AAG ATC AAT GAA GAA Val Asp Leu Val Ala Val Asn Thr Asn Leu Leu Lys Ile Asn Glu Glu 485 490 495	1544
TTA GTT CAA CAA GAG CAA GTG CTC TTA TCG ATG ATT GAC AAT TTT GCA Leu Val Gln Gln Glu Gln Val Leu Leu Ser Met Ile Asp Asn Phe Ala 500 505 510 515	1592
GAA AGT GAA GAG AAT CAA GCC TTG ATT GAA TCG ATG CGT CTT CTT TTG Glu Ser Glu Glu Asn Gln Ala Leu Ile Glu Ser Met Arg Leu Leu Leu 520 525 530	1640
AGA GGC GGT CAT GAT GAG TAA AAAGAGTCCA CAATTAAGGT TTGAAGGTTT Arg Gly Gly His Asp Glu *	1691
535	
TACGGATGA	1700

(2) INFORMATIONS POUR LA SEQ ID NO: 8:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 538 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: protéine

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 8:

```

Met Glu Arg Arg Ile Ile Met Ala Thr Gly Leu Asn Gln Gln Leu Trp
 1           5           10           15
Ala Ser Ala Asp Ile Leu Arg Gly Lys Met Asp Ala Ser Glu Tyr Lys
          20           25           30
Asn Tyr Leu Leu Gly Leu Ile Phe Tyr Lys Tyr Leu Ser Asp Ala Gln
          35           40           45
Leu Arg Glu Val Tyr Glu Gln Glu Asn Gly Lys Thr Asp Thr Phe Pro
          50           55           60
Glu Arg Ser Thr Gln Tyr Ala Gly Phe Met Glu Trp Tyr Glu Glu Asp
          65           70           75           80
Lys Asp Asp Leu Ile Glu Asn Ile Gln Pro Lys Gln Gly Tyr Phe Ile
          85           90           95
Gln Pro Asp Gln Leu Phe Tyr Ser Tyr Arg Ile Lys Ala Asp Asn Tyr
          100          105          110
Glu Phe Asn Leu Thr Asp Leu Gln Ala Gly Phe Asn Glu Leu Glu Arg
          115          120          125
Gln Gly Glu Glu Phe Ser Gly Leu Phe Ala Asp Ile Asp Leu Asn Ser
          130          135          140
Thr Lys Leu Gly Ser Asn Ala Leu Leu Arg Asn Val Thr Ile Thr Glu
          145          150          155          160
Val Leu Arg Ala Leu Asp Glu Ile Asp Leu Phe Glu His Asn Gly Asp
          165          170          175
Val Ile Gly Asp Ala Tyr Glu Tyr Leu Ile Gly Glu Phe Ala Ser Ser
          180          185          190
-----
Ala Gly Lys Lys Ala Gly Glu Phe Tyr Thr Pro Gln Ala Val Ser Lys
          195          200          205
Ile Met Ser Glu Ile Thr Ser Ile Gly Gln Glu Thr Arg Ala Pro Phe
          210          215          220
His Ile Tyr Asp Pro Ala Met Gly Ser Gly Ser Leu Met Leu Asn Ile
          225          230          235          240
Arg Arg Tyr Leu Asn Asn Pro Asp Gln Val His Tyr His Gly Gln Glu
          245          250          255
Leu Asn Thr Thr Thr Phe Asn Leu Ala Arg Met Asn Leu Ile Leu His
          260          265          270

```

Gly Ile Asp Lys Glu Arg Met Asn Leu Asn Asn Gly Asp Thr Leu Asp
 275 280 285
 Ala Asp Trp Pro Ser Glu Glu Pro Tyr Gln Phe Asp Ser Val Cys Met
 290 295 300
 Asn Pro Pro Tyr Ser Ala Lys Trp Ser Ala Ala Asp Gln Phe Leu Ser
 305 310 315 320
 Asp Pro Arg Phe Glu Arg Phe Gly Lys Leu Ala Pro Lys Ser Lys Ala
 325 330 335
 Asp Phe Ala Phe Leu Leu His Gly Phe Tyr His Leu Lys Glu Ser Gly
 340 345 350
 Thr Met Gly Ile Val Leu Pro His Gly Val Leu Phe Arg Gly Ala Ala
 355 360 365
 Glu Gly Thr Ile Arg Gln Ala Leu Leu Glu Met Gly Ala Ile Asp Ala
 370 375 380
 Val Ile Gly Leu Pro Ala Asn Ile Phe Phe Gly Thr Ser Ile Pro Thr
 385 390 395 400
 Thr Val Ile Ile Leu Lys Arg Asn Arg Ser Arg Arg Asp Val Leu Phe
 405 410 415
 Ile Asp Ala Ser Gln Asp Phe Glu Lys Arg Lys Asn Gln Asn Val Leu
 420 425 430
 Leu Asp Glu His Ile Asp Lys Ile Val Ser Ile His Lys Lys Arg Glu
 435 440 445
 Asp Ile Glu Arg Tyr Ala His Val Ala Ser Phe Asp Glu Ile Gln Glu
 450 455 460
 Asn Asp Phe Asn Leu Asn Ile Pro Arg Tyr Val Asp Thr Phe Glu Glu
 465 470 475 480
 Glu Glu Pro Val Asp Leu Val Ala Val Asn Thr Asn Leu Leu Lys Ile
 485 490 495
 Asn Glu Glu Leu Val Gln Gln Glu Gln Val Leu Leu Ser Met Ile Asp
 500 505 510
 Asn Phe Ala Glu Ser Glu Glu Asn Gln Ala Leu Ile Glu Ser Met Arg
 515 520 525
 Leu Leu Leu Arg Gly Gly His Asp Glu *
 530 535

(2) INFORMATIONS POUR LA SEQ ID NO: 9:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1400 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ix) CARACTERISTIQUE:

(A) NOM/CLE: CDS

(B) EMPLACEMENT: 15..1385

(D) AUTRES INFORMATIONS: /product= "HsdS"

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 9:

CTGGGAGTGC TAAA ATG AAG GAA AGA TTA AAG GCT CCA GAA TTA AGG TTT	50
Met Lys Glu Arg Leu Lys Ala Pro Glu Leu Arg Phe	
1 5 10	
GAT GGT TTT ACG GAT GAT TGG GAA GAG CGT AAG TTA TTA GAC AAT GTA	98
Asp Gly Phe Thr Asp Asp Trp Glu Glu Arg Lys Leu Leu Asp Asn Val	
15 20 25	
GAA AAA GTA TTA CAT TAT CGT GGT AAA AGC CCC GCT AAG TTT GGA ATG	146
Glu Lys Val Leu His Tyr Arg Gly Lys Ser Pro Ala Lys Phe Gly Met	
30 35 40	
GAA TGG GGT ACA GAA GGT TAT CTT GTT CTT TCT GCT TTG AAT GTT AAA	194
Glu Trp Gly Thr Glu Gly Tyr Leu Val Leu Ser Ala Leu Asn Val Lys	
45 50 55 60	
AAT GGA TAC ATT GAT AAA TCT GTT GAA GCA AAG TAT GGG GAT CAT GAA	242
Asn Gly Tyr Ile Asp Lys Ser Val Glu Ala Lys Tyr Gly Asp His Glu	
65 70 75	
TTA TTT GAC AGA TGG ATG GGG AAC AAT CGG TTA GAA AAA GGA GAT GTT	290
Leu Phe Asp Arg Trp Met Gly Asn Asn Arg Leu Glu Lys Gly Asp Val	
80 85 90	
GTT TTT ACA ACA GAA GCT CCT TTA GGC AAT GTA GCT CAA GTA CCC GAT	338
Val Phe Thr Thr Glu Ala Pro Leu Gly Asn Val Ala Gln Val Pro Asp	
95 100 105	
AAT AAC GGA TAT ATA TTA AAC CAA AGA GCG GTT GCA TTT AAA TCT TTG	386
Asn Asn Gly Tyr Ile Leu Asn Gln Arg Ala Val Ala Phe Lys Ser Leu	
110 115 120	
CAA GAA ACA GAT GAT AAT TTC TTT GCG CAA TTA CTA CGA AGT CCT ATT	434
Gln Glu Thr Asp Asp Asn Phe Phe Ala Gln Leu Leu Arg Ser Pro Ile	
125 130 135 140	
-----GTT CAA AAT ACA TTA AAA GCT AGT TCT TCT GGG GGA ACT GCC AAA GGT-----	482
Val Gln Asn Thr Leu Lys Ala Ser Ser Ser Gly Gly Thr Ala Lys Gly	
145 150 155	
ATT GGA ATG AAA GAA TTT GCT AAA CTT AAT GCT CGA GTT CCC GAA ACA	530
Ile Gly Met Lys Glu Phe Ala Lys Leu Asn Ala Arg Val Pro Glu Thr	
160 165 170	
CAT GAA GAA CAA CGA AAA ATA GGG TTA TTC TTC AAA CAG CTA GAC GAC	578
His Glu Glu Gln Arg Lys Ile Gly Leu Phe Phe Lys Gln Leu Asp Asp	
175 180 185	
ACT ATC GTT CTT CAT CAA CGT AAG TTA GAT CTT CTC AAA GAG CAG AAA	626
Thr Ile Val Leu His Gln Arg Lys Leu Asp Leu Lys Glu Gln Lys	
190 195 200	

AAA GGA TAC TTG CAA AAA ATG TTC CCC AAA AAT GGT TCA AAA ATT CCT	674
Lys Gly Tyr Leu Gln Lys Met Phe Pro Lys Asn Gly Ser Lys Ile Pro	
205 210 215 220	
GAA TTG AGA TTT GCG GAG TTT GCT GAC GAT TGG GAA GAA CGT AAG TTG	722
Glu Leu Arg Phe Ala Glu Phe Ala Asp Asp Trp Glu Glu Arg Lys Leu	
225 230 235	
GGT GAA GTT GCA ACA TTT CTA AAC GGA CGA GCA TAT AAG CAA GAT GAA	770
Gly Glu Val Ala Thr Phe Leu Asn Gly Arg Ala Tyr Lys Gln Asp Glu	
240 245 250	
TTA CTG GAT TCA GGT AAA TAT AAA GTG CTT CGC GTC GGC AAT TTT TAT	818
Leu Leu Asp Ser Gly Lys Tyr Lys Val Leu Arg Val Gly Asn Phe Tyr	
255 260 265	
ACC AAT GAT TCT TGG TAC TAC TCT AAT ATG GAG CTT GGT GAT AAA TAT	866
Thr Asn Asp Ser Trp Tyr Tyr Ser Asn Met Glu Leu Gly Asp Lys Tyr	
270 275 280	
TAT GTT GAT AAA GGT GAT CTA GTT TAT ACT TGG TCA GCG ACA TTT GGC	914
Tyr Val Asp Lys Gly Asp Leu Val Tyr Thr Trp Ser Ala Thr Phe Gly	
285 290 295 300	
CCA CAT ATT TGG AGT GGC GAA AAA GTT ATT TAC CAT TAT CAC ATC TGG	962
Pro His Ile Trp Ser Gly Glu Lys Val Ile Tyr His Tyr His Ile Trp	
305 310 315	
AAA GTC GAG CTT TCT AAA TTT CTT GAT AGG AAT TTT ACA TTG CAA CTT	1010
Lys Val Glu Leu Ser Lys Phe Leu Asp Arg Asn Phe Thr Leu Gln Leu	
320 325 330	
TTA GAA GCG GAT AAA GCA AGA TTG TTA TCC AGT ACA AAT GGT TCA ACA	1058
Leu Glu Ala Asp Lys Ala Arg Leu Leu Ser Ser Thr Asn Gly Ser Thr	
335 340 345	
ATG ATT CAT GTA ACC AAG GGA GAT ATG GAA AGT AAA ATT GTT TCT ATT	1106
Met Ile His Val Thr Lys Gly Asp Met Glu Ser Lys Ile Val Ser Ile	
350 355 360	
CCT AAT ATT GAT GAG CAA AAA CAA ATT GGT TCA TTC TTC AAA CAA CTC	1154
Pro Asn Ile Asp Glu Gln Lys Gln Ile Gly Ser Phe Phe Lys Gln Leu	
365 370 375 380	
GAC AAC ACT ATC ACC CTT CAT CAA CGT AAG TTA GAT TTG TTG AAA GAG	1202
Asp Asn Thr Ile Thr Leu His Gln Arg Lys Leu Asp Leu Leu Lys Glu	
385 390 395	
CAG AAA AAA GGA TTT CTA CAA AAA ATG TTT CAT TTA ACA AAT TTA GGA	1250
Gln Lys Lys Gly Phe Leu Gln Lys Met Phe His Leu Thr Asn Leu Gly	
400 405 410	
GCC TTC ACT ATG GAA ACT ATA CAA AAA TAT TTA AAG ATA ATT GTT GAA	1298
Ala Phe Thr Met Glu Thr Ile Gln Lys Tyr Leu Lys Ile Ile Val Glu	
415 420 425	

AGT AAT TAT CTT CAT CAA CGT AAG TTA GAC TTG CTT AAG GAG CAA AAA 1346
 Ser Asn Tyr Leu His Gln Arg Lys Leu Asp Leu Leu Lys Glu Gln Lys
 430 435 440

AGA ATT ATT ACA AAT AAT GTT TAT TCA TTT TTT GTA TAA TATAAAGCGA 1395
 Arg Ile Ile Thr Asn Asn Val Tyr Ser Phe Phe Val *
 445 450 455

TTTAA 1400

(2) INFORMATIONS POUR LA SEQ ID NO: 10:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 457 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: protéine

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 10:

Met Lys Glu Arg Leu Lys Ala Pro Glu Leu Arg Phe Asp Gly Phe Thr
 1 5 10 15

Asp Asp Trp Glu Glu Arg Lys Leu Leu Asp Asn Val Glu Lys Val Leu
 20 25 30

His Tyr Arg Gly Lys Ser Pro Ala Lys Phe Gly Met Glu Trp Gly Thr
 35 40 45

Glu Gly Tyr Leu Val Leu Ser Ala Leu Asn Val Lys Asn Gly Tyr Ile
 50 55 60

Asp Lys Ser Val Glu Ala Lys Tyr Gly Asp His Glu Leu Phe Asp Arg
 65 70 75 80

Trp Met Gly Asn Asn Arg Leu Glu Lys Gly Asp Val Val Phe Thr Thr
 85 90 95

Glu Ala Pro Leu Gly Asn Val Ala Gln Val Pro Asp Asn Asn Gly Tyr
 100 105 110

Ile Leu Asn Gln Arg Ala Val Ala Phe Lys Ser Leu Gln Glu Thr Asp
 115 120 125

Asp Asn Phe Phe Ala Gln Leu Leu Arg Ser Pro Ile Val Gln Asn Thr
 130 135 140

Leu Lys Ala Ser Ser Ser Gly Gly Thr Ala Lys Gly Ile Gly Met Lys
 145 150 155 160

Glu Phe Ala Lys Leu Asn Ala Arg Val Pro Glu Thr His Glu Glu Gln
 165 170 175

Arg Lys Ile Gly Leu Phe Phe Lys Gln Leu Asp Asp Thr Ile Val Leu
 180 185 190

His Gln Arg Lys Leu Asp Leu Leu Lys Glu Gln Lys Lys Gly Tyr Leu
 195 200 205
 Gln Lys Met Phe Pro Lys Asn Gly Ser Lys Ile Pro Glu Leu Arg Phe
 210 215 220
 Ala Glu Phe Ala Asp Asp Trp Glu Glu Arg Lys Leu Gly Glu Val Ala
 225 230 235 240
 Thr Phe Leu Asn Gly Arg Ala Tyr Lys Gln Asp Glu Leu Leu Asp Ser
 245 250 255
 Gly Lys Tyr Lys Val Leu Arg Val Gly Asn Phe Tyr Thr Asn Asp Ser
 260 265 270
 Trp Tyr Tyr Ser Asn Met Glu Leu Gly Asp Lys Tyr Tyr Val Asp Lys
 275 280 285
 Gly Asp Leu Val Tyr Thr Trp Ser Ala Thr Phe Gly Pro His Ile Trp
 290 295 300
 Ser Gly Glu Lys Val Ile Tyr His Tyr His Ile Trp Lys Val Glu Leu
 305 310 315 320
 Ser Lys Phe Leu Asp Arg Asn Phe Thr Leu Gln Leu Leu Glu Ala Asp
 325 330 335
 Lys Ala Arg Leu Leu Ser Ser Thr Asn Gly Ser Thr Met Ile His Val
 340 345 350
 Thr Lys Gly Asp Met Glu Ser Lys Ile Val Ser Ile Pro Asn Ile Asp
 355 360 365
 Glu Gln Lys Gln Ile Gly Ser Phe Phe Lys Gln Leu Asp Asn Thr Ile
 370 375 380
 Thr Leu His Gln Arg Lys Leu Asp Leu Leu Lys Glu Gln Lys Lys Gly
 385 390 395 400
 Phe Leu Gln Lys Met Phe His Leu Thr Asn Leu Gly Ala Phe Thr Met
 405 410 415
 Glu Thr Ile Gln Lys Tyr Leu Lys Ile Ile Val Glu Ser Asn Tyr Leu
 420 425 430
 His Gln Arg Lys Leu Asp Leu Leu Lys Glu Gln Lys Arg Ile Ile Thr
 435 440 445
 Asn Asn Val Tyr Ser Phe Phe Val *
 450 455

(2) INFORMATIONS POUR LA SEQ ID NO: 11:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1308 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: double
- (D) CONFIGURATION: linéaire

(ix) CARACTERISTIQUE:

(A) NOM/CLE: CDS

(B) EMLACEMENT: 44..1258

(D) AUTRES INFORMATIONS: /product= "HsdS"

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 11:

GCCTTGATTG AATCGATGCG TCTTCTTTTG AGAGGCGGTC ATG ATG AGT AAA AAG	55
Met Met Ser Lys Lys	
1 5	
AGT CCA CAA TTA AGG TTT GAA GGT TTT ACG GAT GAT TGG GAA GAA CGT	103
Ser Pro Gln Leu Arg Phe Glu Gly Phe Thr Asp Asp Trp Glu Glu Arg	
10 15 20	
AAG TTT GGA GAA GTT TGG AAA AAA TCA AGT GAG CGC AAT TTA AAT TTA	151
Lys Phe Gly Glu Val Trp Lys Lys Ser Ser Glu Arg Asn Leu Asn Leu	
25 30 35	
GAA TAT TCA CCG AAG CAA GTT TTA TCA GTG GCT CAA ATG AAA TTA AAT	199
Glu Tyr Ser Pro Lys Gln Val Leu Ser Val Ala Gln Met Lys Leu Asn	
40 45 50	
CCA TCT AAT AGA AAT GAA CAG GAT GAT TAT ATG AAA ACA TAT AAT GTG	247
Pro Ser Asn Arg Asn Glu Gln Asp Asp Tyr Met Lys Thr Tyr Asn Val	
55 60 65	
CTT CAT AAA GGT GAT ATT GCA TTT GAA GGT AAC AAG TCA AAG AGT TTT	295
Leu His Lys Gly Asp Ile Ala Phe Glu Gly Asn Lys Ser Lys Ser Phe	
70 75 80 85	
GCA TTT GGT AGG TTT GTT CTA GAT GAT TTA CAG GAT GGG ATA GTC TCG	343
Ala Phe Gly Arg Phe Val Leu Asp Asp Leu Gln Asp Gly Ile Val Ser	
90 95 100	
CAT GTA TTT TAT GTA TAT CGC CCT ATT TGT AAA ATG GAT ACA GAT TTT	391
His Val Phe Tyr Val Tyr Arg Pro Ile Cys Lys Met Asp Thr Asp Phe	
105 110 115	
ATG ATA GTT TAT ATA AAT AAT GAA TCT GTA ATG AAG TAT CTT TTA GTT	439
Met Ile Val Tyr Ile Asn Asn Glu Ser Val Met Lys Tyr Leu Leu Val	
120 125 130	
AAG GCA ACT ACT AAA ACT TTG ATG ATG ACT ACT TTA AAT ACT AAA GAT	487
Lys Ala Thr Thr Lys Thr Leu Met Met Thr Thr Leu Asn Thr Lys Asp	
135 140 145	
ATT GTT AAA CCA AAA CTA AAC TTA CCT AGT CTT GAA GAA CAA CAA AAG	535
Ile Val Lys Pro Lys Leu Asn Leu Pro Ser Leu Glu Glu Gln Gln Lys	
150 155 160 165	
ATC GGT TCA TTC TTC AAA CAG TTA GAT GCC ACT ATC GCT CTT CAT CAA	583
Ile Gly Ser Phe Phe Lys Gln Leu Asp Ala Thr Ile Ala Leu His Gln	
170 175 180	
CGT AAG CTA GAT TTG TTG AAA GAA CAG AAA AAA GGC TAC TTC CAA AAA	631
Arg Lys Leu Asp Leu Leu Lys Glu Gln Lys Lys Gly Tyr Phe Gln Lys	
185 190 195	

ATG TTC CCT AAA AAT GGT GCC AAA GTT CCT GAA TTG CGA TTT GCG GGG Met Phe Pro Lys Asn Gly Ala Lys Val Pro Glu Leu Arg Phe Ala Gly 200 205 210	679
TTT GCT GAC GAT TGG GAA GAT CGT AAG TTA GGT GAA TTA GCT AGT TTT Phe Ala Asp Asp Trp Glu Asp Arg Lys Leu Gly Glu Leu Ala Ser Phe 215 220 225	727
TCA AAA GGC AAT GGA TAT ACA AAA AAT GAT TTA GTC GAA TTT GGA GAT Ser Lys Gly Asn Gly Tyr Thr Lys Asn Asp Leu Val Glu Phe Gly Asp 230 235 240 245	775
CCA ATA ATT TTA TAT GGT CGC TTA TAT ACA AAA TAT GAA ACA GTT ATT Pro Ile Ile Leu Tyr Gly Arg Leu Tyr Thr Lys Tyr Glu Thr Val Ile 250 255 260	823
GAG AAA GTA GAT ACA TTT GTA AAT AAG AAA GAT AAA TCA ATA ATT AGT Glu Lys Val Asp Thr Phe Val Asn Lys Lys Asp Lys Ser Ile Ile Ser 265 270 275	871
GGA GGA TCT GAA GTC ATA GTG CCT GCA TCA GGA GAA TCT TCA GAA GAT Gly Gly Ser Glu Val Ile Val Pro Ala Ser Gly Glu Ser Ser Glu Asp 280 285 290	919
ATT TCT AGA GCT TCA GTT GTT GGG AAA TCA GGT ATA ATT TTA GGC GGA Ile Ser Arg Ala Ser Val Val Gly Lys Ser Gly Ile Ile Leu Gly Gly 295 300 305	967
GAT TTA AAT ATA ATT AAA CCA GTA AAT TAT ATT GAT TCT ATT TTT TTA Asp Leu Asn Ile Ile Lys Pro Val Asn Tyr Ile Asp Ser Ile Phe Leu 310 315 320 325	1015
GCT TTA ACT ATT TCT AAT GGA TCT CAA CAA AAA GAA ATG TCT AAG AGA Ala Leu Thr Ile Ser Asn Gly Ser Gln Gln Lys Glu Met Ser Lys Arg 330 335 340	1063
GCT CAA GGA AAA TCT GTA GTT CAT CTG CAT AAC TCT GAT CTA AAA CAA Ala Gln Gly Lys Ser Val Val His Leu His Asn Ser Asp Leu Lys Gln 345 350 355	1111
GTA AAT ATT TTA TAT CCA AAA TTA GGA GAA CAA CAA AAA ATC GGT TCA Val Asn Ile Leu Tyr Pro Lys Leu Gly Glu Gln Gln Lys Ile Gly Ser 360 365 370	1159
TTC TTC AAA CAA CTA GAT AAC ACT ATC GTT CTT CAT CAA CGT AAG TTA Phe Phe Lys Gln Leu Asp Asn Thr Ile Val Leu His Gln Arg Lys Leu 375 380 385	1207
GAT TTT TTG AAA GAG CAG AAA AAA GGC TTT TTA CAA AAG ATG TTT GTT Asp Phe Leu Lys Glu Gln Lys Lys Gly Phe Leu Gln Lys Met Phe Val 390 395 400 405	1255
TAG GGTCTATAAT TAGATAATAA CCCCTCACAA ATCCCATTAA AATAGCCCCT *	1308

(2) INFORMATIONS POUR LA SEQ ID NO: 12:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 406 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: protéine

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 12:

Met	Met	Ser	Lys	Lys	Ser	Pro	Gln	Leu	Arg	Phe	Glu	Gly	Phe	Thr	Asp	1	5	10	15
Asp	Trp	Glu	Glu	Arg	Lys	Phe	Gly	Glu	Val	Trp	Lys	Lys	Ser	Ser	Glu	20	25	30	
Arg	Asn	Leu	Asn	Leu	Glu	Tyr	Ser	Pro	Lys	Gln	Val	Leu	Ser	Val	Ala	35	40	45	
Gln	Met	Lys	Leu	Asn	Pro	Ser	Asn	Arg	Asn	Glu	Gln	Asp	Asp	Tyr	Met	50	55	60	
Lys	Thr	Tyr	Asn	Val	Leu	His	Lys	Gly	Asp	Ile	Ala	Phe	Glu	Gly	Asn	65	70	75	80
Lys	Ser	Lys	Ser	Phe	Ala	Phe	Gly	Arg	Phe	Val	Leu	Asp	Asp	Leu	Gln	85	90	95	
Asp	Gly	Ile	Val	Ser	His	Val	Phe	Tyr	Val	Tyr	Arg	Pro	Ile	Cys	Lys	100	105	110	
Met	Asp	Thr	Asp	Phe	Met	Ile	Val	Tyr	Ile	Asn	Asn	Glu	Ser	Val	Met	115	120	125	
Lys	Tyr	Leu	Leu	Val	Lys	Ala	Thr	Thr	Lys	Thr	Leu	Met	Met	Thr	Thr	130	135	140	
Leu	Asn	Thr	Lys	Asp	Ile	Val	Lys	Pro	Lys	Leu	Asn	Leu	Pro	Ser	Leu	145	150	155	160
Glu	Glu	Gln	Gln	Lys	Ile	Gly	Ser	Phe	Phe	Lys	Gln	Leu	Asp	Ala	Thr	165	170	175	
Ile	Ala	Leu	His	Gln	Arg	Lys	Leu	Asp	Leu	Leu	Lys	Glu	Gln	Lys	Lys	180	185	190	

Gly	Tyr	Phe	Gln	Lys	Met	Phe	Pro	Lys	Asn	Gly	Ala	Lys	Val	Pro	Glu	195	200	205	
Leu	Arg	Phe	Ala	Gly	Phe	Ala	Asp	Asp	Trp	Glu	Asp	Arg	Lys	Leu	Gly	210	215	220	
Glu	Leu	Ala	Ser	Phe	Ser	Lys	Gly	Asn	Gly	Tyr	Thr	Lys	Asn	Asp	Leu	225	230	235	240
Val	Glu	Phe	Gly	Asp	Pro	Ile	Ile	Leu	Tyr	Gly	Arg	Leu	Tyr	Thr	Lys	245	250	255	
Tyr	Glu	Thr	Val	Ile	Glu	Lys	Val	Asp	Thr	Phe	Val	Asn	Lys	Lys	Asp	260	265	270	

Lys Ser Ile Ile Ser Gly Gly Ser Glu Val Ile Val Pro Ala Ser Gly
 275 280 285
 Glu Ser Ser Glu Asp Ile Ser Arg Ala Ser Val Val Gly Lys Ser Gly
 290 295 300
 Ile Ile Leu Gly Gly Asp Leu Asn Ile Ile Lys Pro Val Asn Tyr Ile
 305 310 315 320
 Asp Ser Ile Phe Leu Ala Leu Thr Ile Ser Asn Gly Ser Gln Gln Lys
 325 330 335
 Glu Met Ser Lys Arg Ala Gln Gly Lys Ser Val Val His Leu His Asn
 340 345 350
 Ser Asp Leu Lys Gln Val Asn Ile Leu Tyr Pro Lys Leu Gly Glu Gln
 355 360 365
 Gln Lys Ile Gly Ser Phe Phe Lys Gln Leu Asp Asn Thr Ile Val Leu
 370 375 380
 His Gln Arg Lys Leu Asp Phe Leu Lys Glu Gln Lys Lys Gly Phe Leu
 385 390 395 400
 Gln Lys Met Phe Val *
 405

(2) INFORMATIONS POUR LA SEQ ID NO: 13:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1300 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: double
- (D) CONFIGURATION: linéaire

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: CDS
- (B) EMBLACEMENT: 11..1288
- (D) AUTRES INFORMATIONS: /product= "Hsds"

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 13:

GGTGGGGACT ATG GCG AAA ATA GAT GAT TCA GTT AAA AAG AGA GTT CCA 49
 Met Ala Lys Ile Asp Asp Ser Val Lys Lys Arg Val Pro
 1 5 10
 GAA TTA AGG TTT CCG GGA TTT ACG AAT GAT TGG GAA GAG CGT AAG TTT 97
 Glu Leu Arg Phe Pro Gly Phe Thr Asn Asp Trp Glu Glu Arg Lys Phe
 15 20 25
 TTT GAA AGT ATA GCT TCA ACA ATA GAT TTT AGA GGT AGA ACT CCT AAA 145
 Phe Glu Ser Ile Ala Ser Thr Ile Asp Phe Arg Gly Arg Thr Pro Lys
 30 35 40 45
 AAG TTA GGC ATG GAC TGG AGT GAT TCT GGA TAT TTA GCT TTA TCC GCT 193
 Lys Leu Gly Met Asp Trp Ser Asp Ser Gly Tyr Leu Ala Leu Ser Ala
 50 55 60

TTG AAC GTA AAA AAT GGA TAT ATT GAT CCG TTA GCT GAT GCT CAC TAT	241
Leu Asn Val Lys Asn Gly Tyr Ile Asp Pro Leu Ala Asp Ala His Tyr	
65 70 75	
GGT GAT GAG AAA TTA TAC AGA AAA TGG ATG TCA GGA AGA GAA CTA AAA	289
Gly Asp Glu Lys Leu Tyr Arg Lys Trp Met Ser Gly Arg Glu Leu Lys	
80 85 90	
AAA GGA CAG GTT CTC TTT ACA ACA GAG GCT CCT ATG GGA AAT GTC GCT	337
Lys Gly Gln Val Leu Phe Thr Glu Ala Pro Met Gly Asn Val Ala	
95 100 105	
CAA GTA CCT GAT GAT AAC GGA TAT ATT TTA AGT CAA AGA ACT GTA GCA	385
Gln Val Pro Asp Asp Asn Gly Tyr Ile Leu Ser Gln Arg Thr Val Ala	
110 115 120 125	
TTT GAG ACC AAA GAA GAT ATG ATG ACT AAT GAC TTC TTA GCT GTA TTA	433
Phe Glu Thr Lys Glu Asp Met Met Thr Asn Asp Phe Leu Ala Val Leu	
130 135 140	
TTA AAA TCT CCT TTA GTT TTC AAT AAT TTA TCA GCA TTA TCA AGC GGT	481
Leu Lys Ser Pro Leu Val Phe Asn Asn Leu Ser Ala Leu Ser Ser Gly	
145 150 155	
GGA ACT GCT AAA GGG GTT AGT CAA AAA TCA TTA AAA GGA CTA TCT ATA	529
Gly Thr Ala Lys Gly Val Ser Gln Lys Ser Leu Lys Gly Leu Ser Ile	
160 165 170	
ACT GTC CCA TTG GAT ATT GAC GAG CAA CAA AAA ATC GGT TCA TTC TTC	577
Thr Val Pro Leu Asp Ile Asp Glu Gln Gln Lys Ile Gly Ser Phe Phe	
175 180 185	
AAA CAT TTA GAT GAC ACT ATC GCT CTT CAT CAA CGT AAG TTG GAT TTA	625
Lys His Leu Asp Asp Thr Ile Ala Leu His Gln Arg Lys Leu Asp Leu	
190 195 200 205	
CTC AAG GAA CAG AAA AAA GGC TAC TTG CAA AAA ATG TTC CCT AAA AAT	673
Leu Lys Glu Gln Lys Lys Gly Tyr Leu Gln Lys Met Phe Pro Lys Asn	
210 215 220	
GGT GCC AAA GTT CCT GAA TTG CGA TTT GCG GGG TTT GCT GAC GAT TGG	721
Gly Ala Lys Val Pro Glu Leu Arg Phe Ala Gly Phe Ala Asp Asp Trp	
-----225-----230-----235-----	
GAA GAG CGT AAG TTG GGT GAT ATA GCT CCC TTA CGT GGC GGT TAT GCA	769
Glu Glu Arg Lys Leu Gly Asp Ile Ala Pro Leu Arg Gly Gly Tyr Ala	
240 245 250	
TTC AAA AGC TCT AAA TTC AGG AAA ACT GGT GTG CCA ATT GTT AGA ATT	817
Phe Lys Ser Ser Lys Phe Arg Lys Thr Gly Val Pro Ile Val Arg Ile	
255 260 265	
TCA AAT ATT CTT TCC AGC GGG GAG GTA GGT GGA GAT TTT GCA TAT TAT	865
Ser Asn Ile Leu Ser Ser Gly Glu Val Gly Gly Asp Phe Ala Tyr Tyr	
270 275 280 285	

GAT GAA CAG GAT AAG GAT GAC AAA TAT ATT CTT CCA GAT AAA TCA GCA Asp Glu Gln Asp Lys Asp Asp Lys Tyr Ile Leu Pro Asp Lys Ser Ala 290 295 300	913°
GTC CTA GCC ATG TCA GGT GCA ACA ACA GGT AAG GTA TCT ATA CTA TCT Val Leu Ala Met Ser Gly Ala Thr Thr Gly Lys Val Ser Ile Leu Ser 305 310 315	961
CAA ACT GAT TAT GAC AAG GTC TAT CAA AAC CAG CGA GTT GGT TAT TTT Gln Thr Asp Tyr Asp Lys Val Tyr Gln Asn Gln Arg Val Gly Tyr Phe 320 325 330	1009
CAG TCT GTA GAC TAT ATT GAC TAC GGG TTT ATT TCC ACA ATC GTC CGT Gln Ser Val Asp Tyr Ile Asp Tyr Gly Phe Ile Ser Thr Ile Val Arg 335 340 345	1057
TCA GAA TTA TTC ATG ATG CAA CTT GAG TCT GTT CTA GTT TCA GGC GCT Ser Glu Leu Phe Met Met Gln Leu Glu Ser Val Leu Val Ser Gly Ala 350 355 360 365	1105
CAG CCA AAT GTA TCG TCA AAA GAA ATT GAT TCG TTT AAT TTT ATG ATT Gln Pro Asn Val Ser Ser Lys Glu Ile Asp Ser Phe Asn Phe Met Ile 370 375 380	1153
CCC ATA TTA GTT CAG GAG CAA CAA AAA ATC GGT TCA TTC TTC AAA CAG Pro Ile Leu Val Gln Glu Gln Gln Lys Ile Gly Ser Phe Phe Lys Gln 385 390 395	1201
TTA GAT GAC ACT ATC GCT CTT CAT CAA CGT AAG CTA GAT TTG TTG AAA Leu Asp Asp Thr Ile Ala Leu His Gln Arg Lys Leu Asp Leu Leu Lys 400 405 410	1249
GAA CAG AAA AAA GGC TTT TTA CAA AAG ATG TTT GTT TAG GGTCTATAAT Glu Gln Lys Lys Gly Phe Leu Gln Lys Met Phe Val *	1298
TA	1300

(2) INFORMATIONS POUR LA SEQ ID NO: 14:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 426 acides aminés
- (B) TYPE: acide aminé
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: protéine

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 14:

Met	Ala	Lys	Ile	Asp	Asp	Ser	Val	Lys	Lys	Arg	Val	Pro	Glu	Leu	Arg
1								5			10			15	
Phe	Pro	Gly	Phe	Thr	Asn	Asp	Trp	Glu	Glu	Arg	Lys	Phe	Phe	Glu	Ser
			20					25						30	
Ile	Ala	Ser	Thr	Ile	Asp	Phe	Arg	Gly	Arg	Thr	Pro	Lys	Lys	Leu	Gly
								35						40	45

Met Asp Trp Ser Asp Ser Gly Tyr Leu Ala Leu Ser Ala Leu Asn Val
 50 55 60
 Lys Asn Gly Tyr Ile Asp Pro Leu Ala Asp Ala His Tyr Gly Asp Glu
 65 70 75 80
 Lys Leu Tyr Arg Lys Trp Met Ser Gly Arg Glu Leu Lys Lys Gly Gln
 85 90 95
 Val Leu Phe Thr Thr Glu Ala Pro Met Gly Asn Val Ala Gln Val Pro
 100 105 110
 Asp Asp Asn Gly Tyr Ile Leu Ser Gln Arg Thr Val Ala Phe Glu Thr
 115 120 125
 Lys Glu Asp Met Met Thr Asn Asp Phe Leu Ala Val Leu Leu Lys Ser
 130 135 140
 Pro Leu Val Phe Asn Asn Leu Ser Ala Leu Ser Ser Gly Gly Thr Ala
 145 150 155 160
 Lys Gly Val Ser Gln Lys Ser Leu Lys Gly Leu Ser Ile Thr Val Pro
 165 170 175
 Leu Asp Ile Asp Glu Gln Gln Lys Ile Gly Ser Phe Phe Lys His Leu
 180 185 190
 Asp Asp Thr Ile Ala Leu His Gln Arg Lys Leu Asp Leu Leu Lys Glu
 195 200 205
 Gln Lys Lys Gly Tyr Leu Gln Lys Met Phe Pro Lys Asn Gly Ala Lys
 210 215 220
 Val Pro Glu Leu Arg Phe Ala Gly Phe Ala Asp Asp Trp Glu Glu Arg
 225 230 235 240
 Lys Leu Gly Asp Ile Ala Pro Leu Arg Gly Gly Tyr Ala Phe Lys Ser
 245 250 255
 Ser Lys Phe Arg Lys Thr Gly Val Pro Ile Val Arg Ile Ser Asn Ile
 260 265 270
 Leu Ser Ser Gly Glu Val Gly Gly Asp Phe Ala Tyr Tyr Asp Glu Gln
 275 280 285
 Asp Lys Asp Asp Lys Tyr Ile Leu Pro Asp Lys Ser Ala Val Leu Ala
 290 295 300
 Met Ser Gly Ala Thr Thr Gly Lys Val Ser Ile Leu Ser Gln Thr Asp
 305 310 315 320
 Tyr Asp Lys Val Tyr Gln Asn Gln Arg Val Gly Tyr Phe Gln Ser Val
 325 330 335
 Asp Tyr Ile Asp Tyr Gly Phe Ile Ser Thr Ile Val Arg Ser Glu Leu
 340 345 350

Phe Met Met Gln Leu Glu Ser Val Leu Val Ser Gly Ala Gln Pro Asn
 355 360 365

Val Ser Ser Lys Glu Ile Asp Ser Phe Asn Phe Met Ile Pro Ile Leu
 370 375 380

Val Gln Glu Gln Gln Lys Ile Gly Ser Phe Phe Lys Gln Leu Asp Asp
 385 390 395 400

Thr Ile Ala Leu His Gln Arg Lys Leu Asp Leu Leu Lys Glu Gln Lys
 405 410 415

Lys Gly Phe Leu Gln Lys Met Phe Val *
 420 425

(2) INFORMATIONS POUR LA SEQ ID NO: 15:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1250 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: double
- (D) CONFIGURATION: linéaire

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: CDS
- (B) EMBLACEMENT: 46..1233
- (D) AUTRES INFORMATIONS: /product= "HsdS"

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 15:

AAAAAGGATA GAAGAACAAG AAAAACCAAG AAAGTGGTGG GGACT ATG GCG AAA	54
Met Ala Lys	
1	
ATA GAT GAT TCA GTT AAA AAG AAA GTT CCT GAA TTG CGA TTT AAA GGA	102
Ile Asp Asp Ser Val Lys Lys Lys Val Pro Glu Leu Arg Phe Lys Gly	
5 10 15	
TTT ACG AAT GAT TGG GAA GAG CGT AAG TTA GGA GAA TTA TCT AAT ATT	150
Phe Thr Asn Asp Trp Glu Glu Arg Lys Leu Gly Glu Leu Ser Asn Ile	
20 25 30 35	
GTT GGT GGT GGA ACA CCA AGT ACA TCG AAC CCT GAA TAC TGG GAC GGT	198
Val Gly Gly Gly Thr Pro Ser Thr Ser Asn Pro Glu Tyr Trp Asp Gly	
40 45 50	
GAT ATT GAT TGG TAT GCA CCA GCT GAA ATT GGA GAA CAA AGT TAT GTT	246
Asp Ile Asp Trp Tyr Ala Pro Ala Glu Ile Gly Glu Gln Ser Tyr Val	
55 60 65	
AGC AAA AGT AAA AAG ACT ATT ACT GAA CTA GGT CTA AAG AAG AGT TCA	294
Ser Lys Ser Lys Lys Thr Ile Thr Glu Leu Gly Leu Lys Lys Ser Ser	
70 75 80	
GCT CGA ATT TTA CCA GTA GGA ACC GTC TTA TTT ACT TCT CGT GCT GGT	342
Ala Arg Ile Leu Pro Val Gly Thr Val Leu Phe Thr Ser Arg Ala Gly	
85 90 95	

ATC GGA AAC ACC GCT ATA TTA GCC AAA GAA GCT ACA ACT AAT CAA GGG Ile Gly Asn Thr Ala Ile Leu Ala Lys Glu Ala Thr Thr Asn Gln Gly 100 105 110 115	390
TTT CAA TCA ATT GTT CCT GAT CAA AAT AAA CTT GAT AGT TAT TTT ATT Phe Gln Ser Ile Val Pro Asp Gln Asn Lys Leu Asp Ser Tyr Phe Ile 120 125 130	438
TTT TCA AGA ACT AAT GAG CTC AAG CGA TAT GGG GAA GTC ACT GGT GCA Phe Ser Arg Thr Asn Glu Leu Lys Arg Tyr Gly Glu Val Thr Gly Ala 135 140 145	486
GGA TCT ACT TTT GTT GAG GTT TCA GGT AAG CAA ATG TCG AAA ATG TCG Gly Ser Thr Phe Val Glu Val Ser Gly Lys Gln Met Ser Lys Met Ser 150 155 160	534
ATT ATG GTA CCT GAA CTT TCA GAA CAA CAA AAA ATT GGT AAC TTC TTT Ile Met Val Pro Glu Leu Ser Glu Gln Gln Lys Ile Gly Asn Phe Phe 165 170 175	582
AAA GAG TTA GAC AAC ACT ATC GCT CTT CAT CAG CGT AAG TTA GAT TTG Lys Glu Leu Asp Asn Thr Ile Ala Leu His Gln Arg Lys Leu Asp Leu 180 185 190 195	630
TTG AAA GAA CAG AAA AAA GGC TAC TTG CAA AAA ATG TTC CCT AAA AAT Leu Lys Glu Gln Lys Lys Gly Tyr Leu Gln Lys Met Phe Pro Lys Asn 200 205 210	678
GGT GCC AAA GTT CCT GAA TTG CGA TTT GCG GGG TTT GCT GAC GAT TGG Gly Ala Lys Val Pro Glu Leu Arg Phe Ala Gly Phe Ala Asp Asp Trp 215 220 225	726
GAA GAG CGT AAG TTA GGA GAT ATT ACT AAA ATA AGC ACC GGA AAA TTA Glu Glu Arg Lys Leu Gly Asp Ile Thr Lys Ile Ser Thr Gly Lys Leu 230 235 240	774
GAT GCA AAT GCA ATG GTT GAA AAT GGC AAG TAT GAT TTT TAT ACT TCT Asp Ala Asn Ala Met Val Glu Asn Gly Lys Tyr Asp Phe Tyr Thr Ser 245 250 255	822
GGA ATT AAA AAA TAT AGG ATC GAT GTA GCA GCG TTT GAA GGA CCG TCC Gly Ile Lys Lys Tyr Arg Ile Asp Val Ala Ala Phe Glu Gly Pro Ser 260 265 270 275	870
ATT ACA ATT GCA GGT AAT GGA GCA ACA GTT GGA TAC ATG CAT TTA GCG Ile Thr Ile Ala Gly Asn Gly Ala Thr Val Gly Tyr Met His Leu Ala 280 285 290	918
GAT AAT AAG TTC AAT GCC TAT CAA CGA ACT TAT GTG CTC CAA GAG TTT Asp Asn Lys Phe Asn Ala Tyr Gln Arg Thr Tyr Val Leu Gln Glu Phe 295 300 305	966
CTA GTA GAT AGA AGT TTT ATA TTT TCA GAA ATT GGA AAT AAG CTG CCT Leu Val Asp Arg Ser Phe Ile Phe Ser Glu Ile Gly Asn Lys Leu Pro 310 315 320	1014

AAA AAA ATA AAA CAG GAA GCT CGT ACA GGA AAT ATT CCG TAC ATT GTT	1062
Lys Lys Ile Lys Gln Glu Ala Arg Thr Gly Asn Ile Pro Tyr Ile Val	
325 330 335	
ATG GAC ATG TTG ACT GAA CTG AAG CTA TCG ATT CCT CAG AAT AAT TCA	1110
Met Asp Met Leu Thr Glu Leu Lys Leu Ser Ile Pro Gln Asn Asn Ser	
340 345 350 355	
GAG CAA CAA AAA ATA GGT TCA TTC TTT AAA CAG CTA GAT GAC ACT ATC	1158
Glu Gln Gln Lys Ile Gly Ser Phe Phe Lys Gln Leu Asp Asp Thr Ile	
360 365 370	
GCT CTT CAT CAA CGT AAG TTA GCA TTT GTT GAA AGA ACA GAA AAC ACA	1206
Ala Leu His Gln Arg Lys Leu Ala Phe Val Glu Arg Thr Glu Asn Thr	
375 380 385	
GGC TTT TTA CAA AAG ATG TTT GTT TAG GGTCTATAAT TAGATAA	1250
Gly Phe Leu Gln Lys Met Phe Val *	
390 395	

(2) INFORMATIONS POUR LA SEQ ID NO: 16:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 396 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: protéine

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 16:

Met Ala Lys Ile Asp Asp Ser Val Lys Lys Lys Val Pro Glu Leu Arg	
1 5 10 15	
Phe Lys Gly Phe Thr Asn Asp Trp Glu Glu Arg Lys Leu Gly Glu Leu	
20 25 30	
Ser Asn Ile Val Gly Gly Gly Thr Pro Ser Thr Ser Asn Pro Glu Tyr	
35 40 45	
Trp Asp Gly Asp Ile Asp Trp Tyr Ala Pro Ala Glu Ile Gly Glu Gln	
50 55 60	
Ser Tyr Val Ser Lys Ser Lys Lys Thr Ile Thr Glu Leu Gly Leu Lys	
65 70 75 80	
Lys Ser Ser Ala Arg Ile Leu Pro Val Gly Thr Val Leu Phe Thr Ser	
85 90 95	
Arg Ala Gly Ile Gly Asn Thr Ala Ile Leu Ala Lys Glu Ala Thr Thr	
100 105 110	
Asn Gln Gly Phe Gln Ser Ile Val Pro Asp Gln Asn Lys Leu Asp Ser	
115 120 125	
Tyr Phe Ile Phe Ser Arg Thr Asn Glu Leu Lys Arg Tyr Gly Glu Val	
130 135 140	

Thr Gly Ala Gly Ser Thr Phe Val Glu Val Ser Gly Lys Gln Met Ser																	
145					150				155								160
Lys Met Ser Ile Met Val Pro Glu Leu Ser Glu Gln Gln Lys Ile Gly																	
				165				170									175
Asn Phe Phe Lys Glu Leu Asp Asn Thr Ile Ala Leu His Gln Arg Lys																	
			180					185									190
Leu Asp Leu Leu Lys Glu Gln Lys Lys Gly Tyr Leu Gln Lys Met Phe																	
			195				200										205
Pro Lys Asn Gly Ala Lys Val Pro Glu Leu Arg Phe Ala Gly Phe Ala																	
			210				215										220
Asp Asp Trp Glu Glu Arg Lys Leu Gly Asp Ile Thr Lys Ile Ser Thr																	
225					230					235							240
Gly Lys Leu Asp Ala Asn Ala Met Val Glu Asn Gly Lys Tyr Asp Phe																	
				245					250								255
Tyr Thr Ser Gly Ile Lys Lys Tyr Arg Ile Asp Val Ala Ala Phe Glu																	
			260					265									270
Gly Pro Ser Ile Thr Ile Ala Gly Asn Gly Ala Thr Val Gly Tyr Met																	
			275					280									285
His Leu Ala Asp Asn Lys Phe Asn Ala Tyr Gln Arg Thr Tyr Val Leu																	
			290					295									300
Gln Glu Phe Leu Val Asp Arg Ser Phe Ile Phe Ser Glu Ile Gly Asn																	
305					310					315							320
Lys Leu Pro Lys Lys Ile Lys Gln Glu Ala Arg Thr Gly Asn Ile Pro																	
				325					330								335
Tyr Ile Val Met Asp Met Leu Thr Glu Leu Lys Leu Ser Ile Pro Gln																	
			340					345									350
Asn Asn Ser Glu Gln Gln Lys Ile Gly Ser Phe Phe Lys Gln Leu Asp																	
			355					360									365
Asp Thr Ile Ala Leu His Gln Arg Lys Leu Ala Phe Val Glu Arg Thr																	
370							375										380
Glu Asn Thr Gly Phe Leu Gln Lys Met Phe Val *																	
385							390										395

REVENDEICATIONS

1) Polypeptide, constituant l'une des sous-unités HsdR, HsdM, ou HsdS d'un mécanisme de résistance aux bactériophages R/M de type Ic, caractérisé en ce que
5 ledit mécanisme est actif contre les phages des bactéries lactiques.

2) Polypeptide selon la revendication 1, caractérisé en ce qu'il est choisi dans le groupe constitué par :

10 - les polypeptides constituant une sous-unité HsdR comprenant la séquence (I) suivante :

KRLARK

- les polypeptides constituant une sous-unité HsdM comprenant au moins l'une des séquences suivantes :

15 - la séquence (II):

GLX₁FYKYLS

dans laquelle X₁ représente I ou L,

- la séquence (III):

ENDX₂NLNIPRYVDTFEEEE

20 dans laquelle X₂ représente Y ou F

- les polypeptides constituant une sous-unité HsdS comprenant :

* un domaine N-terminal d'environ 150 à 180 acides aminés, comprenant la séquence (IV) suivante :

25 PX₃LRFX₄GFTX₅DDWEERKX₆

dans laquelle X₃ représente E ou Q, X₄ représente E, P, D, ou K, X₅ représente N ou D, X₆ représente L ou F ;

* un domaine central, d'environ 50 à 80 acides aminés, comprenant la séquence (V) suivante :

30 BOX₇KIGX₈FFKX₉LDX₁₀TX₁₁LHQRKLDLLKEQKGYX₁₂QKMFPKNGX₁₃KX₁₄PELRFAX₁₅FADWEX₁₆RKLG

dans laquelle X₇ représente R ou Q, X₈ représente S, N, ou L, X₉ représente E, H, ou Q, X₁₀ représente A, D, ou N, X₁₁ représente A, ou V, X₁₂
35 représente L ou F, X₁₃ représente A ou S, X₁₄ représente I ou V, X₁₅ représente E ou G, X₁₆ représente D ou E ;

* un domaine C-terminal, d'environ 160 à 200 acides aminés, comprenant la séquence (VI) suivante :

EQX₁₅X₁₆IGSF₁₇FKQLDX₁₇TIX₁₈LHQ₁₉RKLX₁₉

5 dans laquelle X₁₅ représente K ou Q, X₁₆ représente K ou Q, X₁₇ représente N ou D, X₁₈ représente T, A, ou V, X₁₉ représente A, ou D ;

et/ou la séquence (VII) suivante :

GFLQKMF₂₀X₂₀

dans laquelle X₂₀ représente V ou H.

10 3) Polypeptide selon la revendication 2, caractérisé en ce qu'il est choisi dans le groupe constitué par :

- un polypeptide HsdR répondant à l'une des séquences respectivement représentées dans la liste de
15 séquences en annexe sous les numéros SEQ ID NO:2 et SEQ ID NO:4 ;

- un polypeptide HsdM répondant à l'une des séquences respectivement représentées dans la liste de
20 séquences en annexe sous les numéros SEQ ID NO:6 , et SEQ ID NO:8 ;

- un polypeptide HsdS constitué par :

* un domaine N-terminal d'environ 150 à 180 acides aminés dont la séquence est celle du
25 domaine N-terminal de l'une quelconque des séquences représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16 ;

30 * un domaine central, d'environ 50 à 80 acides aminés, dont la séquence est celle du domaine central de l'une quelconque des séquences représentées dans la liste de séquences en annexe sous les numéros SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16 ;

35 * un domaine C-terminal, d'environ 160 à 200 acides aminés, dont la séquence est celle du

domaine C-terminal de l'une quelconque des
séquences représentées dans la liste de
séquences en annexe sous les numéros
SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14,
5 SEQ ID NO:16.

4) Séquence d'acide nucléique caractérisée en
ce qu'elle code pour un polypeptide selon une quelconque
des revendications 1 à 3.

5) Séquence d'acide nucléique selon la
10 revendication 4 caractérisée en ce qu'elle est choisie
dans le groupe constitué par :

- les séquences *hsdR* SEQ ID NO:1, et SEQ ID NO:3 ;
- les séquences *hsdM* SEQ ID NO:5, et SEQ ID NO:7 ;
- les séquences *hsdS* SEQ ID NO:9, SEQ ID NO:11,
15 SEQ ID NO:13, et SEQ ID NO:15.

6) Fragment d'acide nucléique d'au moins 18 pb
homologue ou complémentaire de tout ou partie d'une
séquence d'acide nucléique selon une quelconque des
revendications 4 ou 5.

7) Fragment d'acide nucléique selon la
20 revendication 6, choisi dans le groupe constitué par les
oligonucléotides homologues ou complémentaires d'une
séquence codant pour au moins 6 acides aminés consécutifs
de l'un des peptides suivants ;

25 TGSGKT
KRLARK
LLTGFD
FYKYLS
LARMNL
30 LPHGVLFGRGAAE
NLNIPRYVDTFEEEE
DDWEERK
LHQRKLDLLKEQKKGY
QKMFPKNG
35 PELRFA
FADDWE

IGSFFKQLD

GFLQKMF.

5 8) Utilisation d'au moins un fragment d'acide
nucléique selon une quelconque des revendications 6 ou 7
pour sélectionner des bactéries lactiques contenant une
séquence d'acide nucléique codant pour au moins une sous-
unité d'un mécanisme R/M de type Ic.

10 9) Utilisation d'au moins un fragment d'acide
nucléique selon une quelconque des revendications 6 ou 7
pour isoler et/ou cloner un acide nucléique comprenant une
séquence selon une quelconque des revendications 4 ou 5.

15 10) Utilisation d'au moins une séquence
d'acide nucléique selon une quelconque des revendications
4 ou 5 pour permettre l'expression dans une bactérie
lactique d'au moins un mécanisme de résistance aux
bactériophages.

11) Bactérie transformée par au moins une
séquence d'acide nucléique selon une quelconque des
revendications 4 ou 5.

20 12) Bactérie transformée selon la Reven-
dication 11, caractérisée en ce qu'elle est en outre
capable d'exprimer un mécanisme de résistance aux
bactériophages choisi parmi les mécanismes Abi, et les
mécanismes R/M de type II.

INSTITUT NATIONAL

RAPPORT DE RECHERCHE

N° d'enregistrement
national

PRELIMINAIRE

d la
PROPRIETE INDUSTRIELLEétabli sur la base des dernières revendications
déposées avant le commencement de la rechercheFA 547763
FR 9710885

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
D,X	A. CHOPIN ET AL., : "Two plasmid-determined restriction and modification systems in Streptococcus lactis" PLASMID, vol. 11, 1984, NEY YORK, NY, US, pages 260-263, XP002066556 * le document en entier *	1,3-5
Y	---	12
X	M. GAUTIER AND M.C. CHOPIN: "Plasmid-determined systems for restriction and modification activity and abortive infection in Streptococcus cremoris" APPLIED ENVIRONMENTAL MICROBIOLOGY, vol. 53, no. 5, 1987, US, pages 923-927, XP000573466 * le document en entier *	1,3-5
Y	---	12
Y	US 5 629 182 A (M.C. CHOPIN ET AL) 13 mai 1997 * le document en entier *	12
Y	---	12
Y	WO 96 21017 A (QUEST INT) 11 juillet 1996 * le document en entier *	12
X	E. EMOND ET AL., : "Phenotypic and genetic characterization of the bacteriophage abortive infection mechanism AbiK from Lactococcus lactis" APPLIED ENVIRONMENTAL MICROBIOLOGY, vol. 63, no. 4, 1 avril 1997, WASHINGTON, DC,US, pages 1274-1283, XP002067220 * le document en entier *	2,7

-/--		
Date d'achèvement de la recherche		Examineur
25 juin 1998		Mateo Rosell, A.M.
<p>CATEGORIE DES DOCUMENTS CITES</p> <p>X: particulièrement pertinent à lui seul Y: particulièrement pertinent en combinaison avec un autre document de la même catégorie A: pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général O: divulgation non-écrite P: document intercalaire</p> <p>T: théorie ou principe à la base de l'invention E: document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D: cité dans la demande L: cité pour d'autres raisons &: membre de la même famille, document correspondant</p>		

1

EPO FORM 1503 03.82 (P04C13)

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
X	W0 96 25503 A (J. JOSEPHSEN ET AL.,) 22 août 1996 * le document en entier *	12
A	EP 0 452 224 A (SANOFI SA; ELF AQUITAINE) 16 octobre 1991 * le document en entier *	1-12
A	GB 2 294 463 A (UNIV NEW SOUTH WALES ;BURNS PHILIP & CO LTD (AU); BURNS PHILIP RES) 1 mai 1996 * le document en entier *	1-12
A	FR 2 738 015 A (SYSTEMS BIO IND) 28 février 1997 * le document en entier *	1-11
D,A	T.A. BICKLE AND D.H. KRÜGER: "Biology of DNA restriction" MICROBIOLOGICAL REVIEWS, vol. 57, no. 2, - 1993 WASHINGTON, DC, US, pages 434-450, XP002066557 voir le document en entier et tout particulièrement pages 434-439	1-12
		DOMAINES TECHNIQUES RECHERCHES (Int.CL.6)
Date d'achèvement de la recherche		Examineur
25 juin 1998		Mateo Rosell, A.M.
<p>CATEGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant</p>		

1

EPO FORM 150 (3.82 (P44C13))

